

Three Essays on the Law and Economics of Taxation and Finance

by

Edward Gellis Fox

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2018

Doctoral Committee:

Professor James Hines, Jr., Chair
Professor Ian Ayres, Yale University
Professor Paul Rhode
Professor Joel Slemrod

Edward Fox

edfox@umich.edu

ORCID iD: 0000-0002-0662-9569

© Edward Fox 2018

Dedication

For my parents and Nora

Acknowledgements

The following acknowledgments are by their nature incomplete; too many people have contributed directly or indirectly to this dissertation for me to thank them all individually. I am nevertheless indebted to each of them and I try to remember Emerson's instruction to "include all [of them in my] gratitude."

There are nevertheless a few people whom I would be remiss if I did not single out for particular thanks. Jim Hines, the chair of my committee, helped guide me through this process with warmth, wit, and gentle but incisive criticism. He was also always extremely generous with his time. Among my advisors, my debt to Jim is—to borrow from Augustus—first among equals. Paul Rhode kindled my interest in economic history and his encyclopedic knowledge has been both an inspiration and a valuable resource in an advisor. Joel Slemrod's work sparked the third chapter and he provided valuable and forthright critiques on the dissertation. Ron Gilson was an inventive and conscientious coauthor on the second chapter.

I completed a substantial portion of my dissertation while I was a student at Yale Law School and Ian Ayres' advice and creativity encouraged and shaped my work both while I was at Yale and after when he joined my committee. Likewise, I benefited greatly in learning from and working with Alan Schwartz and Henry Hansmann. I would also like to thank David Tabak and Ned Phelps whose support and encouragement was integral in my seeking a Ph.D. in economics in the first place.

Finally, I am indebted most of all to my parents Merritt Fox and Ann Gellis and to my wife Nora Sennett. My parents have been my role models, constant champions, frequent editors

and in the case of my dad, my coauthor. Nora is, in John Donne's words, the "fixed foot" around which I spin. Her love, support, and intelligence, not to mention her patience, has inspired my work and enabled me to complete this dissertation.

Table of Contents

Dedication	ii
Acknowledgements.....	iii
List of Tables	vi
List of Figures	vii
List of Appendices	viii
Abstract	ix
Chapter 1 Introduction	1
Chapter 2 Do Taxes Affect Marriage? Lessons from History	5
Chapter 3 Economic Crisis and the Integration of Law and Finance: The Impact of Volatility Spikes (with Merritt Fox and Ronald Gilson).....	50
Chapter 4 Does Capital Bear the Corporate Tax After All? New Evidence from U.S. Corporate Tax Returns.....	154
References	192
Appendices.....	196

List of Tables

Table 1: Summary Statistics Adult Men in 1947	31
Table 2: Estimated Impact of 1948 Revenue Act on Marriage Rate: Men 18-50	38
Table 3: Impact of 1948 Tax Change on Marriage Rates	39
Table 4: Impact of 1948 Tax Change on Age at First Marriage	41
Table 5: Impact of 1948 Tax Change on Fertility.....	43
Table 6: Idiosyncratic Risk of S&P 100 Firms, 2005–2009	60
Table 7: Corporate Returns Attributable to the Normal Return to Capital Using Gentry and Hubbard Method	170
Table 8: Association between the Portion of the Corporate Tax Raised from Normal Return and the Risk-Free Rate & Intangible Capital.....	190
 Table 1A- 1: Additional Results	 197
 Table 2A- 1: Example of Estimated Idiosyncratic Risk	 200
Table 2A- 2: Idiosyncratic Risk and Co-movement of Returns with the Broad Market	204
Table 2A- 3 Leverage Increase, 2007-2009—Firms in the S&P 100	208
Table 2A- 4: Percentage Change in Risk as a Function of Changes in β	212
 Table 3A- 1: Comparing a 35% Tax on the Normal Rate of Return to Actual Corporate Tax Revenue.....	 218
Table 3A- 2: Cashflows of U.S. Non-Financial Corporations	219
Table 3A- 3: Comparison of Hypothetical R-Base Tax to Actual Corporate Tax 1957-2013 .	220

List of Figures

Figure 1: Map of Community Property States	19
Figure 2: Taxes Owed by a Single-Earner Couple with \$100,000 in Income through Time	26
Figure 3: Marriage Rate among High Income Men.....	32
Figure 4: Marriage Rate of High Income Men - Low Income Men	34
Figure 5: Idiosyncratic Risk 2004–2010.....	59
Figure 6: Idiosyncratic Risk 1925–2010.....	61
Figure 7: Idiosyncratic Risk of S&P 100 Firms Adjusted for Leverage	65
Figure 8: Isolation of Firm-Specific News Through Market-Adjustment.....	77
Figure 9: Demonstration of Type I Error.....	80
Figure 10: Demonstration of Type II Error	84
Figure 11: When the Corrective Disclosure Causes A 5% Drop in Price —Comparing Regular and Volatile Periods.....	117
Figure 12: Visualizing the Reduction in Type II Error in Altering the Threshold	122
Figure 13: Type II Error Reduction from Reducing the Confidence Level for Statistical Significance in Crisis Times versus Normal Times and the Implications for Social Welfare ...	129
Figure 14: Quickturn Idiosyncratic Risk	149
Figure 15: Airgas Idiosyncratic Risk	152
Figure 16: Corporate Taxes on Non-Financial C-Corps vs. Tax on Normal Rate of Return ...	171
Figure 17: Portion of Corporate Tax Lost Switching to Cash Flow Taxation 1995-2013	178
Figure 18: Portion of Corporate Tax Lost Switching to Cash Flow Taxation 1992-2013	181
Figure 19: Gross Fixed Business Investment	189
 Figure 1A- 1: Taxes Owed by A Single-Earner Couple with \$100,000 in Income.....	 196
 Figure 2A- 1: Baseline Estimate Idiosyncratic Risk 1925-2010	 201
Figure 2A- 2: Idiosyncratic Risk 1964-1997—Comparison to Campbell et al	202
Figure 2A- 3: The Effect of Controlling for Industry Returns	203
Figure 2A- 4: The Comparison to Fama-French 3 Factor Model.....	206
Figure 2A- 5: The Effect of Truncating the Tails of Distribution	207
Figure 2A- 6: The Effect of Controlling for Leverage	210
Figure 2A- 7: The Effect of Controlling for Beta Instability with Short Regression Windows	214
Figure 2A- 8: Using Implied Volatilities to Adjust for Beta Instability	215
Figure 2A- 9: OLS vs. GARCH.....	217

List of Appendices

Appendix 1: Appendix to Do Taxes Affect Marriage?.....	196
Appendix 2: Appendix to Economic Crises and the Integration of Law and Finance (with Merritt Fox and Ronald Gilson).....	198
Appendix 3: Appendix to Does Capital Bear the Corporate Tax After All?	218

Abstract

This dissertation empirically measures how the laws governing taxation and finance affect behavior and addresses how those laws should adapt to changing circumstances. The first chapter examines the effect of joint-taxation and “marriage bonuses” on marriage formation in the United States in the late 1940s and early 1950s. It uses a natural experiment to identify the effect and finds that tax incentives caused an increase in the marriage rate of up to 9%. The second chapter shows that idiosyncratic risk has spiked in every economic downturn since the 1920s and develops new models to explain this phenomenon. It then explores the implications of spikes in idiosyncratic risk for corporate and securities law. The third chapter compares the existing corporate tax to a hypothetical “cash flow tax” to determine how much of the corporate tax base is composed of the normal return to capital. It finds that the normal return to capital made up a relatively small percentage of the corporate tax base over the last 20 years. Because taxes on the normal return to capital are the most likely to be passed on to labor, this suggests that labor’s long-run share of the corporate tax burden is likely to be lower than typically thought.

Chapter 1 Introduction

This dissertation contains three additional chapters applying empirical methods to issues in the law of tax and corporations and finance. The second chapter of this dissertation deals with the vexing question of how to tax married couples, which has remained controversial since the adoption of the federal income tax over a century ago. The appropriate answer to this question depends in part on how sensitive couples are to taxes when deciding to marry. Yet we know surprisingly little about how taxes shape couples' marriage choices. The chapter begins to fill that gap using a natural experiment generated by the halting shift in how the income tax treated married couples in the mid-Twentieth Century. The system moved from taxing married couples as two individuals—in which case marriage largely did not affect taxes—to taxing married couples on their joint income. At the time of this shift, joint taxation lowered couples' taxes upon marriage. The change to joint taxation, however, came later to some states, creating a natural experiment to study its impact on marriage rates.

The chapter shows that annual marriage rates increased by 9% in the relevant states after the introduction of joint taxation made marriage tax-advantaged, with affected men marrying 3 to 5 months sooner on average. This suggests that at any given time during this period there were tens or hundreds of thousands of married couples in the United States who would not have been married if not for the tax incentives. Couples appear to have been unexpectedly responsive to the tax changes given that unmarried cohabitation was not acceptable under the social mores of the day. If anything, Americans today are likely more sensitive to taxes when deciding whether and when to marry, suggesting that joint taxation continues to affect marriage decisions today.

This in turn strengthens the case for returning to individual taxation of marriage if the goal is to avoid inefficiently distorting people's marriage decisions. By contrast, if the government wishes to encourage marriage, the results imply that using the tax code may be effective under some circumstances, but further analysis suggests that joint taxation remains a poor choice for doing so.

The third chapter, coauthored with Merritt Fox and Ronald Gilson and published in the *Columbia Law Review*, addresses a puzzle which we first identified during the 2008 recession and important for understanding how the capital market prices common stocks and in turn, for the intersection between law and finance. During the crisis, there was a dramatic five-fold spike, across all industries, in "idiosyncratic risk"—the volatility of individual-firm share prices after adjustment for movements in the market as a whole.

This phenomenon is not limited to the most recent financial crisis. The third chapter uses an empirical review to show that a dramatic spike in idiosyncratic risk has occurred with every major downturn from the 1920s through the recent financial crisis. It canvasses three possible explanations for this phenomenon. Thereafter, the chapter explores the implications of these crisis-induced volatility spikes for certain legal issues that depend analytically on valuation methodology and hence are affected by volatility: using event studies to determine materiality and loss causation in fraud-on-the-market securities litigation, determining materiality in cases involving claims of both insider trading and misstatements or omissions in registered public offerings, and determining the extent of deference given to a corporate board that rejects an acquisition offer at a premium above the pre-offer market price.

This analysis shows that the conventional use of event studies during periods of economic-crisis-induced volatility spikes results in understating the number of occasions when a

corporate misstatement can be shown to have had a meaningful impact on a firm's stock price. Relatedly, the analysis suggests that during crisis times, insiders have substantially more opportunities to profit from trading on the nonpublic information that they possess and issuers conducting offerings have more opportunities to sell securities at an inflated price. Analysis shows that trying to cure this problem by lowering the standard of what is considered statistically significant is as likely to be socially harmful as socially beneficial. These conclusions counsel that the best response to the reduced effectiveness of private litigation as a deterrent to securities law violations during crisis times is to provide additional resources to SEC enforcement. Lastly, with respect to Delaware courts' recognition of "substantive coercion" as a justification for target-corporation deployment of takeover defenses—arguably a dubious justification in normal times—crisis-induced idiosyncratic-risk spikes provide an unusually plausible claim that target shareholders may indeed make a mistake in tendering into a hostile offer. Analysis of the timing of the spikes in recent cases, however, shows that the claim is tenuous even in these circumstances.

The fourth chapter presents new evidence from U.S. corporate tax returns comparing the existing tax to a hypothetical cash flow tax, which exempts the normal return to capital. Under reasonable assumptions, the corporate tax raised only 4% of its revenue from the normal return from 1995 to 2013. The similarity of the revenue raised by the hypothetical cash flow tax and the actual corporate tax has two important implications. First, under the Treasury's model for distributing the corporate tax, 50% of the burden of taxes on the normal return to capital ultimately are borne by labor, while 100% of the burden of taxes on "supernormal returns" like economic rents falls on capital owners. Under this model, these results show that the corporate tax burden falls only on capital owners because the current tax raises essentially no revenue from

the normal return to capital. This result differs from Treasury's current estimate that roughly 40% of the tax is raised from the normal return and thus that 20% of the tax is ultimately borne by labor. Second, the results show that the United States could switch to a cash flow tax without sacrificing much revenue, but the efficiency gains would be smaller than if the current tax raised more revenue from the normal return. Finally, the chapter explores why the tax raised from the normal return appears to have fallen over time, finding both theoretically and empirically that two factors are associated with this decline: (1) decreases in the riskless return to capital as measured by rates on Treasury notes, and (2) increased investment in intangible capital by businesses which has largely been given cash flow treatment under the corporate tax.

Chapter 2 Do Taxes Affect Marriage? Lessons from History

INTRODUCTION

And to get to the point, if we were to file a joint return, it would seem to me that with the two exemptions, a joint return would not only be feasible, but economically sound. Indeed. Taxwise.

Jack Lemmon in *Phffft* (1954) (asking his girlfriend to marry him)

Policymakers face a fundamental choice in implementing a progressive income tax: treat married couples as two individuals, or as a unit. There is an inherent tension in this decision. Treating married couples like any two individuals ensures that their income taxes will not change at marriage or divorce. This means that both marriage decisions and other important choices, like whether to work outside the home, will not be distorted by marriage taxation.¹ Under individual taxation, however, married couples with the same total income may pay different taxes depending on how income is divided among the spouses. In addition, the government must prevent spouses from shifting income so as to reduce their taxes. Taxing couples on their joint income solves these problems, but necessitates having couples' taxes change at marriage. Thus under joint taxation, policymakers must then decide whether couples' taxes will rise or fall at marriage, leading to either "marriage bonuses" or "marriage penalties." These bonuses and particularly penalties are perceived by many as inequitable and the tax changes at marriage may also distort marriage and employment choices.

¹ This study is concerned with income taxes, so when I refer to a "tax" I mean income tax, unless otherwise noted.

There will always be a tension in deciding between individual and joint taxation because the advantages associated with each cannot be fully combined. It is impossible for an income tax to be (1) progressive, (2) marriage neutral, and (3) tax all couples with the equal income the same way (“couples neutrality”).² To see this, consider two couples: W and X, who each earn \$50,000 per year, and Y and Z, who earn \$100,000 and \$0 per year, respectively. Prior to marriage, progressivity implies that Y’s taxes exceed W and X’s combined taxes. Marriage neutrality requires the total taxes of each couple not change at marriage, so Y and Z must pay more tax than W and X after each couple marries. This of course violates couples neutrality. Upholding couples neutrality instead requires either the taxes of W and X or Y and Z change at marriage, or both, violating marriage neutrality. Tax scholars frequently refer to the impossibility of meeting all three goals as the trilemma.³

In theory and practice, progressivity has been seen as too important to give way before the other two principles. From the start of the modern income tax in 1913 to 1948, the United States (largely) chose marriage neutrality by employing individual taxation. Since 1948, it has prioritized couple’s neutrality by using joint taxation, which has often led to large tax changes at marriage. Single-earner couples have consistently received marriage bonuses, with couples’ total taxes falling at marriage. For example, a single-earner couple with income each year

² See, e.g., Tax Treatment of Single Persons and Married Persons Where Both Spouses Are Working: Hearings Before the H. Comm. on Ways and Means, 92d Cong. 78-79 (1972) (statement of Edwin S. Cohen, Ass’t Sec’y for Tax Pol’y).

³ For a sample of the discussion of the appropriate tax treatment of marriage discussing the trilemma, see, e.g., Lily Kahng, *One Is the Loneliest Number: The Single Taxpayer in a Joint Return World*, 61 HASTINGS L.J. 651 (2010); Lawrence Zelenak, *Marriage and the Income Tax*, 67 S. Cal. L. Rev. 339 (1994); Edward McCaffery, *TAXING WOMEN* (1997); Michael J. McIntyre & Oliver Oldman, *Taxation of the Family in a Comprehensive and Simplified Income Tax*, 90 HARV. L. REV. 1572 (1977); Boris I. Bittker, *Federal Income Taxation and the Family*, 27 STAN. L. REV. 1389 (1975). See also Anne Alstott, *Updating the Welfare State: Marriage, the Income Tax, and Social Security in the Age of Individualism*, 66 TAX L. REV. 695 (2014) (arguing that the trilemma should no longer be viewed as central to the taxation of the family given the “deinstitutionalization” of marriage). Yair Listokin, *Taxation and Marriage: A Reappraisal*, 67 TAX L. REV. 185, 191 (2013), also quite rightly points out that although the trilemma is frequently framed as requiring picking *either* couples or marriage neutrality, violating both principles slightly might be better than upholding one neutrality completely while badly violating the other.

equivalent to \$100,000 in today's dollars would have seen their taxes fall on average by \$4,423 *per year* after marriage over the period 1948 to 2016.⁴ By contrast, dual-earner couples have faced marriage penalties on average since a change in law in 1969, when marriage penalties were first made common.⁵ Moreover, although Congress moved to reduce or eliminate marriage penalties as part of the Bush tax cuts,⁶ recent estimates suggest that about 40% of married couples still face marriage penalties, averaging about \$1,500 per year (or \$28 billion total), while 55% enjoyed marriage bonuses averaging about \$2,300 per year (or \$60 billion total).⁷ The remaining roughly 5% of couples saw no change in taxes as a result of their marriage. Even these sizable figures probably understate the size of income tax changes resulting from marriage. The same authors estimate that returning fully to a system of individual taxation would have raised revenue by \$140 billion in 2009, or about 15% of total personal income tax revenue for that year.⁸

Although these tax changes at marriage are longstanding, we know relatively little about how, if at all, they shape marriage decisions. This chapter helps to fill that gap by using a natural

⁴ See *infra*, Appendix, Figure A-1 (assuming the standard deduction is taken by the single-earner prior to marriage and by the couple during marriage and the non-working spouse's personal exemption was used prior to marriage).

⁵ James Alm & J. Sebastian Leguizamon, *Whither The Marriage Tax?*, 68 NAT'L. TAX J. 251, 263 (2015). Marriage penalties were first technically imposed by the creation of head of household status in 1951. See Lawrence Zelenak, *Doing Something About Marriage Penalties: A Guide for the Perplexed*, 54 Tax L. Rev. 1, 70 (2000).

⁶ Pub. L. No. 107-16, §§ 101, 301-302, 115 Stat. 38, 41, 53-54 (codified as amended at I.R.C. §§ 1(f), 1(i), 63(c) (2012)); Lawrence Zelenak, *For Better and Worse: The Differing Income Tax Treatments of Marriage at Difference Income Levels*, 93 N. CAROLINA L. REV. 783, 807 (2014).

⁷ See Alm & Leguizamon, *supra* note 5, at 261. The number of taxpayers facing marriage penalties is arguably inflated (and the number receiving bonuses is therefore likewise arguably too low) because the authors' baseline is the couple's most advantageous tax position if unmarried, including having the higher earner file as head of household if the couple has children. Some people would think of the marriage penalty calculation as being most naturally done relative to a system with only individual filing. This is particularly true because head of household status itself exists (at least originally) only to provide some of the tax benefits of marriage to single parents. See Jacob Goldin & Zachary D. Liscow, *Beyond Head of Household: Rethinking the Taxation of Single Parents*, TAX L. REV. (forthcoming 2017).

⁸ See Alm & Leguizamon, *supra* note 5, at 273 (reform includes eliminating head of household status); OMB, BUDGET OF THE U. S. GOVERNMENT: FISCAL YEAR 2011 at 152 (listing 2009 personal income tax receipts), available at <https://www.gpo.gov/fdsys/pkg/BUDGET-2011-BUD/pdf/BUDGET-2011-BUD.pdf>

experiment generated by what Judge Henry Friendly called marriage taxation's "long and stormy history."⁹ Specifically, under the federal income tax, the shift to joint taxation in the mid-20th century occurred at different times in different states. In that period, joint taxation created only marriage bonuses—whereas marriage had been largely tax neutral before—because it permitted "full income splitting." Under full income splitting, spouses were taxed like two individuals, except that each was attributed 50% of the couple's joint income, regardless of the actual division of income between them. This lowered the taxes of couples with enough income to be taxed progressively.

As a formal matter, the 1948 Revenue Act was to work a nationwide change by adopting joint taxation with full income splitting. In practice, however, the Act's adoption of joint taxation only affected "common law" states, where each spouse owns her wages and income from assets in her name. Joint taxation with full income splitting had already come years before in the other states, which had laws providing that most income received during marriage becomes the "community property" of both spouses. This was because in 1930, the Supreme Court, in *Poe v. Seaborn*,¹⁰ interpreted the tax code to allow each spouse in community property states to report half of the community income on her individual tax return, using the rationale that each had an equal ownership interest in the income. Also important for this study, the 1948 Act had no effect on couples in either type of state whose income was low enough that they did not pay income taxes or else were in the lowest bracket. This is an important additional feature of the natural experiment examined in the chapter because it can help control for confounding

⁹ See *Druker v. Commissioner*, 697 F.2d 46, 48 (2d Cir. 1982) (upholding the constitutionality of marriage penalties).

¹⁰ 282 U.S. 101, 101 (1930).

factors which might affect common law and community property states differently and otherwise be confused for the effect of the Act.

The 1948 Revenue Act sets up a natural “difference-in-difference” design to measure the effect on marriage decisions of the marriage bonus that comes from joint taxation with full income splitting.¹¹ The tax incentives to marry of residents of the community property states and lower income persons in both types of states did not change in 1948 and their behavior can be used as controls for what would have happened among affected individuals in the common law states without the Act. In particular, using the publicly available sample of the 1960 census, I construct a history of first marriages for Americans in the six years before, and six years after, passage of the Act (1942-1947 and 1948-1953).¹² This allows me to calculate the likelihood a person who has never married weds in a given year (hereinafter simply “the marriage rate”). I then examine whether the marriage rate for higher income people in common law states (the “treated group”)¹³ increased in the years following the Act, compared to changes in the control groups.

The evidence shows that after the 1948 Revenue Act was enacted, marriage rates of higher income people in common law states increased, both absolutely and, more importantly,

¹¹ *Seaborn* could also be used as a similar natural experiment, but, as discussed below, only the rich paid income taxes at the time. Only about 3% of American adults filed an income tax return in 1930, compared with about 60% in 1947. See *infra*, note 66. The small population affected by *Seaborn*, as well as data limitations in the 1940 census, limit our ability to study the effect of the case with sufficient precision.

¹² The 1948 Revenue Act was enacted April 2, 1948 after Congress overrode President Truman’s veto and became effective for the entire 1948 tax year. Pub. L. No. 471, 80th Cong., 2d Sess. Truman twice vetoed similar legislation in 1947. See Stanley S. Surrey, *Federal Taxation of the Family: The Revenue Act of 1948*, 61 HARV. L. REV. 1097 (1948). To the extent some couples in common law states married in 1947 in anticipation of joint taxation with income splitting, this will cut against my finding an effect from the 1948 Act because these marriages increase the baseline marriage rate in the period before the Act.

¹³ I will sometimes refer to higher income people in common law states as the “treated or treatment group” to remind the reader as simply as possible whose marriage incentives were changed by the 1948 Act. This language is common in economics and derives from randomized controlled trials in medicine in which there is a treatment and control (placebo) group.

relative to the control groups.¹⁴ For example, at the aggregate level, the marriage rate among treated men increased from an average of 8.9% in the six years before the Act to 10.1% in the six years after (or 1.2%), while the marriage rates of their higher income counterparts in the community property (control) states increased only from 9.7% to 10.2% (or 0.5%). By contrast, marriage rates among low income men showed the opposite pattern. Among those men, marriage rates rose faster in community property states (rising 1.4% compared to 1.0%). This rough cut of the data intuitively suggests that confounding factors were not driving the faster increase in the marriage rate among treated men in common law states because that pattern does not hold among the untreated, lower income men.

In the analysis below, I adjust for the behavior of the control groups and a variety of other individual factors in a difference in difference regression model. I estimate that the marriage rate of treated men in common law states increased by about 9% after the 1948 Act (a 0.8% increase in the rate under the Act compared to a 9.3% baseline marriage rate). The increase in the marriage rate is primarily caused by men marrying 3 to 5 months sooner on average than would have been predicted absent the Act, rather than the Act inducing men who would not otherwise have married to wed. Given that unmarried cohabitation was socially unacceptable at the time—presumably adding to the underlying significance of a couple’s decision to marry—couples appear to have been surprisingly sensitive to tax considerations in making this decision.¹⁵

The Act also appears to have indirectly caused other important changes in couples’ lives, with treated couples having more children. On average 8 in 100 affected couples is estimated to have had an extra child as a result of the Act. This makes sense both because the Act seems to

¹⁴ “Higher” and “lower” income are defined in the Data section, *infra*.

¹⁵ See, e.g., Lawrence Zelenak, *For Better and Worse: The Differing Income Tax Treatments of Marriage at Difference Income Levels*, 93 N. CAROLINA L. REV. 783, 792 (2014) (arguing that when cohabitation was socially unacceptable, we would not expect marriage bonuses to have large effects on marriage decisions).

have caused couples to marry younger and reduced married women's financial incentives to participate in the labor force by effectively increasing their marginal tax rate.¹⁶ Both factors pushed couples toward having more children. Moreover, the 1948 Act increased the take-home income of couples at the beginning of the Baby Boom, which also likely allowed them to have more kids.

Understanding how couples reacted to marriage bonuses in the Post-War period reveals a previously unrecognized role for federal tax policy in increases in marriage and fertility during the Baby Boom. Moreover, it can also shed new light on the long running debate about the appropriate tax treatment of marriage. By maintaining joint taxation since 1948, the United States has broken with a global trend over the latter part of the 20th century of countries moving from joint to individual taxation of marriage. In 1970, just six of twenty-two developed countries used individual taxation, while the rest, like the United States, used joint taxation.¹⁷ From 1970 to 1990, however, ten of those countries, including the United Kingdom and New Zealand, switched to some form of individual taxation.¹⁸ No developed country switched from individual to joint taxation.

¹⁶ During this period, the vast majority of married women earned less than their husbands if they worked outside the home. Under individual taxation, married women would therefore tend to be taxed in lower brackets, but under joint taxation women's marginal income would be taxed at a higher rate because their husband's income would drive up the couple's bracket. *See infra* note 31 and accompanying text.

¹⁷ *See* Congressional Budget Office, *For Better or Worse: Marriage and the Federal Income Tax*, 59 (1997) (discussing joint taxation among members of the Organization for Economic Development and Cooperation (OECD))

¹⁸ Most of these countries adopted individual taxation for wages and salary only, while pooling unearned income among the spouses, and then taxing it either to the higher earner or apportioning it to the spouses by formula. Thus these systems are to some extent a hybrid: individual taxation for wage and salary income and joint taxation for capital income. *See id*; *see also* Oliver Oldman & Ralph Temple, *Comparative Analysis of the Taxation of Married Persons*, 12 STANFORD L. REV. 585, 585-86 (1960) (discussing the tax systems of various countries before the trend towards individual taxation began). One motive, albeit not the main one, for the U.K.'s switch to individual taxation was to avoid discouraging marriage. *See* Stephanie Hunter McMahon, *London Calling: Does the UK's Experience with Individual Taxation Clash with the US's Expectations*, 55 ST. LOUIS U.L.J. 159, 187 (2010).

Scholars have strenuously debated whether the United States should join those other countries.¹⁹ The arguments for choosing either couples neutrality (and joint taxation) or marriage neutrality (and individual taxation) invoke more basic concepts of tax fairness and economic efficiency.²⁰ The results of this chapter strengthen the efficiency case for marriage neutrality, i.e., the libertarian case stemming from the view that the choice to marry should not be distorted by its tax effects. The more sensitive Americans are to taxes in making these decisions, the greater the efficiency cost from joint taxation.²¹ American's unexpected responsiveness to taxes in the Post-War period suggests joint taxation continues to effect marriage decisions today. Indeed, the two most readily identifiable changes in the decades since the study period would push most single-earner couples toward paying more attention to taxes in making marriage decisions. Those changes are (1) the dramatic rise in cohabitation, which has given couples a largely socially acceptable way to live together without marriage, and (2) increases in marriage bonuses today for single-earner couples with income equivalent to the "high income" couples around 1948 (\$50,000 to \$110,000 per year in today's dollars). Thus, although we cannot know with certainty, there are good reasons to believe taxes may exert a larger influence on the marriage decisions of single-earner couples making \$50,000 to \$110,000 today. If true, this would make the estimates in this chapter a floor on the effect of taxes today for those couples.²² To the extent one puts any weight on the desirability of not having taxes

¹⁹ See, *supra* note 3.

²⁰ See Listokin, *supra* note 3, at 191; see also Alm & Leguizamon, *supra* note 5 (noting the efficiency costs of distorting marriage decisions).

²¹ In economic models, it is often assumed that taxes have no impact on the decision to get married. For example, the most prominent theory article on joint taxation makes this assumption. See Henrik Jacobsen Kleven, Claus Thustrup Kreiner & Emmanuel Saez, *The Optimal Income Taxation of Couples*, 77 *ECONOMETRICA* 537 (2009). All modeling requires simplification, but my results suggest that this model omits a reasonably significant feature of the optimal tax problem. I thank Zach Liscow for pointing this out.

²² Note that single-earner couples today constitute a much smaller proportion of married couples than during the study period. On the other hand, today a significantly larger percentage of single-earner couples earn enough

directly shape whether a couple marries or not, the results of the chapter therefore augment the case for returning to individual taxation of married couples.

To the extent the government instead wishes to promote marriage, the results imply that using the tax code can be effective in doing so. Nevertheless, joint taxation with full income splitting for all married taxpayers remains a poor tool for encouraging marriage relative to providing a credit of fixed size to all married taxpayers. The vast majority of the benefits from joint taxation with full income splitting accrue to the richest taxpayers, whom the evidence indicates are less sensitive, per dollar, to tax incentives to marry. Moreover, dual income couples, who now make up a substantial majority of all couples, receive little or no benefits. And joint taxation with income splitting causes a host of other problems. For example, it reinforces gender roles by discouraging married women, where they are the lower earning partner in a heterosexual relationship, from working outside the home and this distortion is also quite inefficient economically.²³

The remainder of this chapter is divided into six parts. The first reviews the literature on how the income tax affects marriage decisions. The second provides background on community property and gives the history of how community property income was taxed by the federal government, leading to *Seaborn* in 1930 and eventually the 1948 Act. The third discusses the census data used in this study. The fourth offers the primary results and analyzes the robustness of the result. The fifth discusses the main results and their implications for policy. The sixth briefly concludes.

income to obtain sizable marriage bonuses. This means that the proportion of all married couples receiving large marriage bonuses is roughly similar today as during the study period. *See, infra* note 113 and accompanying text.

²³ *See, e.g., Zelenak, supra* note 3, at 365.

LITERATURE

This chapter contributes generally to a rich literature empirically estimating how economic factors affect marriage decisions, and more directly to a smaller literature on the role of U.S. taxes. The earliest attempt to empirically identify the impact of U.S. income taxes used aggregate data on the U.S. marriage rate and the size of the marriage penalties and bonuses through time.²⁴ It found a relationship between taxes and marriage, but the reliability of such aggregate analysis is low because it cannot control for unobserved underlying variables that over time correlate with changes in tax penalties/bonuses and also affect marriage rates. For example, the loosening of social attitudes toward marriage starting in the 1960s correlates with increases in marriage penalties through time, making it difficult to ascertain how important these penalties were in explaining the drop in marriage rates. Later papers are more convincing as they tried to estimate this relationship using data at the individual level. The papers which most convincingly identify a causal relationship between marriage and taxes showed that “notches” in the tax code (or transfer system) have an impact on the timing of marriage.²⁵ These notches are places where, by slightly adjusting timing of marriage, taxpayers can significantly change their taxes (or benefits). These studies are useful in showing that couples are paying attention to taxes when deciding when to marry within a short time frame, but the importance of these timing games is usually limited.

²⁴ See, e.g., James Alm & Leslie A. Whittington, *Income Taxes and the Marriage Decision*, 27 APPLIED ECON. 25 (1995). For summaries of the literature, see James Alm, Stacy Dickert-Conlin & Leslie A. Whittington, *Policy Watch: The Marriage Penalty*, 13 J. ECON. PERSPECTIVES 251 (1999); Nancy Burstein, *Economic Influences On Marriage And Divorce*, 26 J. POLICY ANALYSIS AND MGM'T 387 (2007).

²⁵ See James Alm & Leslie A. Whittington, *Income taxes and the Timing of Marital Decisions*, 64 J. PUBLIC ECON. 219 (1997) (finding that marriage penalties often cause couples to delay marriage until the next year (e.g., from November to January) to avoid being taxed as though they were married for the entire year); Michael J. Brien, Stacy Dickert-Conlin & David A. Weaver, *Widows Waiting to Wed?(Re) Marriage and Economic Incentives in Social Security Widow Benefits*, 39 J. Human Resources 585 (2004) (Widows lose claims on their deceased husbands' social security spousal benefits if they remarry before 60 and the authors find this has a significant effect on the timing of remarriage around women's 60th birthday.).

Scholars have also used data on individual marriage decisions to study the broader effect of marriage penalties and bonuses.²⁶ This work finds that income tax incentives affect marriage rates, although the effects are of moderate size. For example, James Alm and Leslie Whittington estimated that increasing marriage penalties by 10% led to a 13% fall in the likelihood of first marriage in a year for women near the mean of their sample from 1968 to 1992. These studies are still hampered, however, by the lack of a natural experiment which makes it difficult to disentangle whether the estimates actually reflect the causal effect of income tax incentives or other factors.

Similarly, a number of studies have examined the effect of transfer programs embedded in the tax code (e.g., the Earned Income Tax Credit) and those outside it (e.g., welfare programs like AFDC and later TANF). Lower income people may face proportionally quite large marriage penalties (and sometimes bonuses) from the ways these programs treat the marriage of potential participants. Most of these studies find the programs affect marriage decisions, but do so in relatively modest ways.²⁷

Non-tax economic factors have been shown to substantially affect marriage rates and so care must be taken to be sure any apparent relation between tax changes and marriage rates is not really the result of these other economic causes. Mathew Hill, using data similar to mine, has shown that marriage rates fell in areas hit harder by the Depression, as measured by retail sales

²⁶ See James Alm & Leslie A. Whittington, *For Love or Money? The Impact of Income Taxes on Marriage*, 66 *ECONOMICA* 297 (1999); see also Hector Chade & Gustavo Ventura, *Income Taxation And Marital Decisions*, 8 *REV. ECON. DYNAMICS* 251 (2005) (creating an analytical model of how marriage decisions are made and calibrating that model using data).

²⁷ See Burstein, *supra* note 24, at 411-418 (summarizing the literature on the effect of AFDC/TANF, the EITC, food stamps, etc). The behavioral effects of these programs on marriage may be blunted in part by how difficult it is for participants (and researchers) to understand “byzantine interactions of the[m]” and what the ultimate effect of getting married will be on take home income. See *id.*

activity. Hill attributes this to male unemployment and lack of disposable income.²⁸ A number of other studies also document the effect of wages and employment on marriage in the United States and elsewhere.²⁹ Therefore, I control for economic conditions at the state level in the analysis below. In addition, I estimate a “fixed effects model” that will control for the effect of any other relevant economic factors to the extent the factors do not change over the time period being studied.³⁰

Prior literature also empirically confirms that the coming of joint taxation also altered work incentives. From 1930 to 1948, a working woman in a community property state would usually see her marginal tax rate increase if she married a high earning man, because her tax rate was driven up by her husband’s income. Theory suggests this would reduce her incentives to continue working outside the home. In common law states during this period, where each member of a couple was taxed individually, this was not true; there taxation of a wife’s income would start at the lowest bracket, no matter her husband’s income. As noted, the 1948 Revenue Act extended joint taxation with income splitting to the common law states. Thus, in these states,

²⁸ See Matthew Hill, *Love in the Time of Depression: The Effect of Economic Downturns on the Probability of Marriage*, 75 J. Econ. History 163 (2015). See also Namkee Ahn & Pedro Mira, *Job Bust, Baby Bust?: Evidence from Spain*, 14 J. POPULATION ECON. 505 (2001) (finding male unemployment decreases the probability of marriage); Humphrey Southall & David Gilbert, *A Good Time to Wed?: Marriage and Economic Distress in England and Wales, 1839-1914*, 49 ECON. HIST. REV. 35 (1996).

²⁹ The literature is substantial. A few papers on the U.S. include: Michael C. Keely, *The Economics of Family Formation*, 15 ECON. INQUIRY 238 (1977) (finding all else equal in the 1960s men with higher wages married slightly sooner, while the opposite is true for women); Ted Bergstrom & Robert F. Schoeni, *Income Prospects and Age-at-Marriage*, 9 J. POPULATION ECON. 115 (1996) (evidence that the effect Keely finds is confined to men marrying under age 30, and the opposite effect for those who have not married by 30); Megan M. Sweeney & Maria Cancian, *The Changing Importance of White Women's Economic Prospects for Assortative Mating*, 66 J. MARRIAGE & FAMILY 1015 (2004) (showing an increasing correlation between a wife’s pre-marriage income and her husband’s income for U.S. marriages).

³⁰ Intuitively, fixed effects allow one to compare only similar individuals in the model. For example, I use state, age, and income group fixed effects, which means that my estimates implicitly only compare persons from the same state, of the same age group, and income group. Mechanically this is implemented by (in essence) including a separate “indicator” variable for each possible combination of state, age, and income group in the regressions below. Many factors which affect the marriage rate, say of high income men in Michigan aged 20-24, will presumably not change significantly over the twelve year period under study. Using fixed effects, those factors will be controlled for, even if they cannot be directly observed and will be represented in the indicator variable associated with that state, age, and income group combination.

suddenly, a wife earning less than her husband saw her marginal tax rate driven up by her husband's income. Using a study design similar to this one, Sara LaLumia examined the impact of the 1948 changes on labor force participation by women in higher income households in common law states, with similar women in community property states as a control.³¹ She found that the 1948 changes reduced women's participation in the workforce in common law states.³²

This study contributes to the literature along a couple of dimensions. First, the chapter examines a clean natural experiment with multiple control groups. This provides an unusually good opportunity to understand the effect of joint taxation with full income splitting on marriage during the Post War period and to be more confident that the estimates are not driven by confounding factors. The results suggest the role of federal tax policy in increases in marriage during this period (and in the Baby Boom) is greater than has been previously understood. This is a matter of intrinsic historic interest.

Second, the results may be informative about the effect of federal tax policy today on marriage decisions. Qualitatively, the finding that couples were surprisingly sensitive to tax considerations during the Post War period, suggests that joint taxation continues to influence marriage decisions today. This is particularly true because the rise of cohabitation has very likely made couples more sensitive to tax considerations in marrying by permitting them a socially acceptable option to live together without marriage. Moreover, for single-earner couples

³¹See Sara LaLumia, *The Effects of Joint Taxation of Married Couples on Labor Supply and Non-Wage Income*, 92 J. PUB. ECON. 1698 (2008).

³² It is worth noting that LaLumia's work shows that income splitting affects marriage incentives for at least two reasons. The first is the "bonus" that a married couple will reap if they do not change their work choices after the new tax system comes into place. One can (roughly) imagine this as the effect of keeping the unit of taxation on the individual, but having the government providing a credit equal to the amount the couple's taxes would decline under joint taxation if they do not change their labor decisions. Second, joint taxation will affect the menu of after tax wages a married couple may choose from usually by decreasing the husband's marginal rate and increasing the wife's. Some couples may find this new menu of after tax wages more appealing than the old menu and that could influence marriage incentives as well (although others, of course, could find the new menu less attractive). In addition, even the effect of the "bonus" itself can be decomposed into an income and substitution effect.

today making \$50,000 to \$110,000—an amount equivalent to the earnings of the high income couples examined in this study around 1948—marriage bonuses are larger now than in 1948.

Quantitatively, inferring lessons for today from the results of this chapter is more complicated and involves a tradeoff: the natural experiment provides clarity arguably unavailable using modern data, but requires one to make strong assumptions about how couples today compare to those many decades ago. While there have been a myriad of changes since the study period which are potentially relevant, the importance of the rise of cohabitation and the large increases in marriage bonuses for single-earner couples making \$50,000 to \$110,000 suggests that the effects estimated here might reasonably be seen as a floor on the effect of taxes today for those couples.

LEGAL BACKGROUND AND HISTORY

A. Community Property

Two marital property regimes prevail in the United States: community property and separate property, the latter being followed in common law states.³³ These regimes are quite different. In common law states, income earned by each spouse remains her own separate property. The same holds for income derived from property legally owned by one spouse, even if the property was bought with funds from a joint bank account.³⁴ In contrast, in community

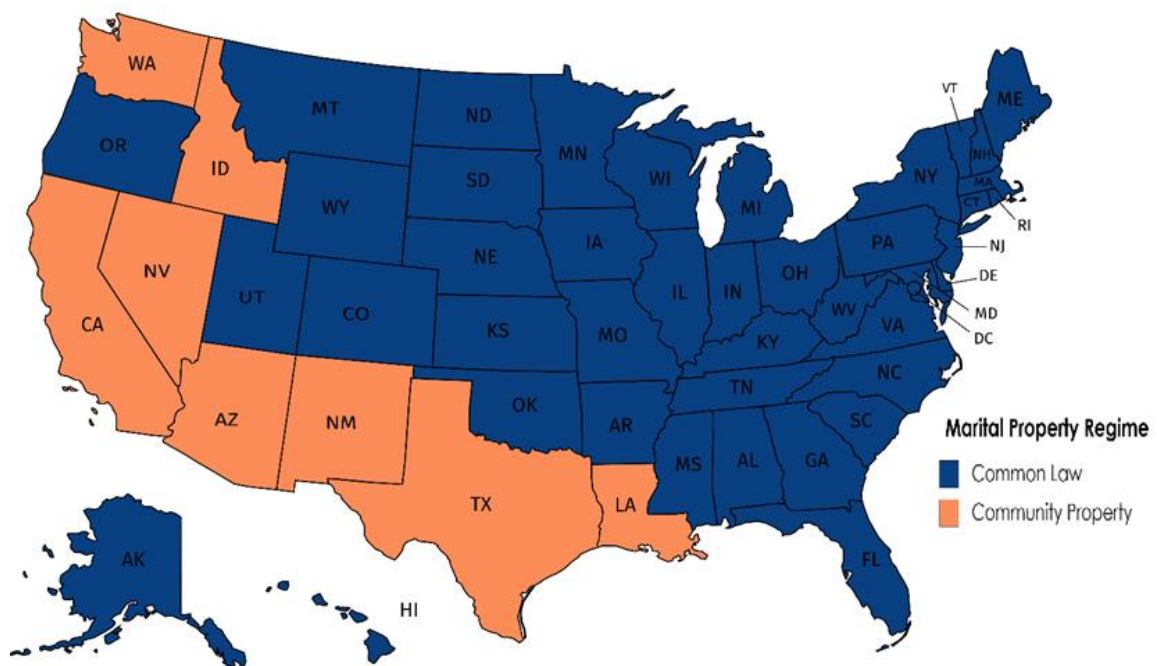
³³ For basic background see JESSE DUKEMINIER & ROBERT SITKOFF, *WILLS, TRUSTS, AND ESTATES* 512-514 (9th ed. 2013). To avoid confusion I will use “common law states” rather than separate property to refer to this regime because separate property has its own meaning in the community property system.

³⁴ Spouses can enter into various forms of joint ownership, e.g. joint tenancy, if they so choose. Since the 1960s even in common law states legal ownership plays a smaller role in the division of property after a divorce under “equitable distribution” statutes. Such statutes tend to give the judge wide latitude to divide the assets regardless of the nominal owner. During the period in question, though, community property states differed significantly from common law states in how property was divided at divorce. For example, *Newsweek* reported that “[i]n [common law] states, if a husband left his wife or if a wife went home to mother, the best she could hope for was a nominal support allowance. In any community property state, the woman automatically came out with half the family bankroll.” *Community Dilemma*, Oct. 13, 1947, at 64-65.

property states “spouses retain separate ownership of property brought to the marriage, but they own all earnings and acquisitions from earnings during the marriage in equal, undivided shares.”³⁵

States have different marital property systems because of their differing legal origins. Community property exists in states with marriage law based in French and Spanish settlement, while common law states derive their marital rules from England. In 1930 at the time of *Seaborn*, Arizona, California, Idaho, Louisiana, Nevada, New Mexico, Texas, and Washington a community property system. These states are highlighted in orange in the map below, while the common law states are in blue.³⁶

Figure 1: Map of Community Property States



³⁵ DUKEMINIER & SITKOFF, *supra* note 33, at 512. It should be noted that within the community property states, laws vary some about narrower questions like how to characterize income received during marriage from separate property.

³⁶ In the 1980s Wisconsin adopted community property, but for all periods relevant to this paper it was a common law state. See Palma Maria Forte, *Wisconsin Marital Property Act: Sections in Need of Reform*, 79 MARQ. L. REV. 859 (1995). As discussed below, five states and Hawaii, which was then a territory, passed community property laws shortly before 1948 to try to obtain the tax advantages under *Seaborn*.

B. History of the Taxation of Marriage and the 1948 Revenue Act

The modern³⁷ federal income tax began its life “dominated by an individualistic approach” to family taxation.³⁸ Each spouse was taxed on the income she or he earned. This largely ensured marriage neutrality.³⁹ However, marriage neutrality also meant that in a progressive tax system, two married couples with the same total income could pay very different taxes. Two-earner couples would pay less tax than single-earner couples with the same income.

In 1913, Congress passed the first modern income tax under the newly approved 16th Amendment. The statute purported to tax “net income arising or accruing from all sources...to every citizen of the United States.”⁴⁰ The 1916 Revenue Act changed the language to tax “the net income of every individual.”⁴¹ This language “of every individual” persisted in each revenue act enacted until *Seaborn* in 1930.⁴²

The question of how for tax purposes to assign income to spouses in community property states proved thorny from the start. Until 1919 the Treasury held that the husband was taxable on all community income.⁴³ In 1919 the Treasury adopted a compromise position. Income from

³⁷ The United States also had an income tax during the Civil War (though it collected little revenue) and 1895, but the latter tax had to be abandoned after the Supreme Court ruled large parts of it unconstitutional in *Pollock v. Farmers' Loan & Trust Company*, 157 U.S. 429 (1895).

³⁸ Bittker, *supra* note 3, at 1400.

³⁹ There were in fact modest marriage bonuses built into the structure of personal exemptions for much of the period before WWII, with personal exemptions for married couples usually more than twice as large as those for single individuals. See Larry Zelenak, *A Troubled Relationship from the Start: Marriage and the Income Tax* (on file with author). This should not affect the empirical analysis below, however, as the exemptions applied equally across both types of states.

⁴⁰ Act of Oct.3, 1913, ch. 16, §II(A)(I), 38 Stat. 166.

⁴¹ Act of Sept.8, 1916, ch. 463, §I(A), 39 Stat. 463.

⁴² *Poe v. Seaborn*, 282 U.S. 101, 101 n. 1 (1930). See also Dennis J. Ventry Jr., *Saving Seaborn: Ownership Not Marriage as the Basis of Family Taxation*. 86 IND. L.J. 1459 (2011) for a quite complete history of the lead up to *Seaborn*.

⁴³ See Douglas Blount Maggs, *Community Property and the Federal Income Tax*, 14 CAL. L. REV. 351, 354 (1926). Maggs reports that “[u]ntil the latter part of 1919 none of the Treasury rulings were made public; for that reason references to specific rulings made before that time cannot be given.” *Id.*

what the Treasury thought of as genuinely belonging to the couple as a whole—e.g., bonds purchased after marriage—would be split 50:50 on each return. A higher earning husband’s salary, however, would be taxed to him in community property states despite the fact that the wife had an identical property interest in the salary as in the income from bonds and other capital.⁴⁴ This treatment of a husband’s salary essentially mirrored what was allowed in common law states. In those states, salaries and wages could not be split between the spouses via intra-family gift or agreement,⁴⁵ but income from capital could be reapportioned using intra-spousal gifts.⁴⁶ This equivalence with the tax treatment of couples in common law states was short-lived, however. In 1920 and 1921 the Attorney General concluded that community property states, except California, gave the wife a vested interest in one half of all community income and thus required splitting *all* community income, including wages and salaries.⁴⁷

In 1926 the Supreme Court upheld California’s exclusion in *Robbins v. United States*.⁴⁸ The Court, via Justice Holmes, agreed with the government that California’s law gave the wife a “mere expectancy,” thus validating California’s exclusion from true community property treatment. But the Court went further and concluded that—even if the wife’s interest were instead vested under California law—Congress could tax all community income to the husband

⁴⁴ *Id.* at 355; Bittker, *supra* note 3, at 1406. Along with salaries and wages, income from separate property purchased before the marriage was to be taxed to the spouse that owned it. This interpretation was tenuous in Idaho, Louisiana and Texas because in those states the income from separate property is (usually) community property. Maggs, *supra* note 43, at 353; *see also* IRS, PUBLICATION 555 at 4 n.1 (2014).

⁴⁵ *See* Lucas v. Earl, 281 U.S. 111 (1930), although the law on splitting income through intra-family agreements was relatively unclear at the time Treasury made its 1919 ruling.

⁴⁶ For example, a husband could use his after-tax salary to buy a bond in his wife’s name. Such a gift would make the interest on the bond taxable to the wife rather than the husband. *See* Bittker, *supra* note 3, at 1401 and 1403. La Lumia finds a decline in women in common law states reporting non-wage income following the 1948 Act. This is consistent with couples previously shifting income to avoid taxes. However, even in relatively high income households, 90% of women reported (to the census) that they had less than \$50 of non-wage income, which suggests that the majority of high income couples did not make use of income shifting. La Lumia, *supra* note 31, at 1711.

⁴⁷ *See* 32 Ops. Att’y Gen. 298; *see also* Maggs, *supra* note 43, at 355; Bittker, *supra* note 3, at 1406. In California, the Attorney General found that during her husband’s life “the wife has no vested interest in the community property, her interest therein being a mere expectancy.” 32 Ops. Att’y Gen. 435, 456.

⁴⁸ 269 U.S. 315 (1926).

“if it so minded.” After reviewing the power of the husband to manage the property,⁴⁹ the Court concluded that in fact “it was intended [by Congress] to tax him for the whole.”⁵⁰ Contemporary commentators viewed the opinion as “in effect an invitation (or a command) to the Treasury to reconsider” its decision to allow income splitting in community property states other than California.⁵¹

Following *Robbins*, the Attorney General withdrew his 1920 and 1921 opinions and the Treasury moved to tax all community income to the husband. This led to *Seaborn*, in which a Washington couple challenged the Treasury’s ruling that the husband had to include all community income. There the Supreme Court reversed course and held that all community income should be split by the spouses. The Court found that the “of” in “net income of every individual” must connote ownership⁵² and that under the laws of Washington “the entire property and income of the community can no more be said to be that of the husband, than it could rightly be termed that of the wife.”⁵³ Taking a more formal approach, the Court found that the husband’s broad managerial powers under the family law of the day were not relevant; the wife’s *rights* in the property were equal to his. The opinion all but ignores *Robbins*, dismissing it as specific to California law and ignoring its conclusion that Congress intended to tax all California community income to the husband even if the wife had a vested interest in it.⁵⁴ It is difficult to explain the Court’s seemingly abrupt reversal. It is true that Holmes’ conclusion about

⁴⁹ In Holmes’ usual high rhetoric, the Court found that “he who has all the power [should] bear the burden and [thus]. . . the husband [is] the most obvious target for the shaft.” *Id.* at 328. Moreover, the Court observed that if the wife was unable or unwilling to pay, the government would apparently be unable to seize community property. *Id.* (“[B]ut the [community] fund taxed, while liable to be taken for his debts, is not liable to be taken for the wife’s . . . so that the remedy for her failure to pay might be hard to find.”).

⁵⁰ *Id.* at 327.

⁵¹ Maggs, *supra* note 43, at 362.

⁵² 282 U.S. at 109.

⁵³ *Id.* at 113

⁵⁴ *Id.* at 116.

Congress' intent was dicta, but it is still puzzling that five Justices, including Justice Holmes, joined both *Robbins* and *Seaborn* taking seemingly opposite interpretations of the same statutory language just four years later.⁵⁵

In sister cases released the same day as *Seaborn*, the Court held that community income should also be split in Arizona, Texas, and Louisiana.⁵⁶ The Treasury quickly allowed the same treatment to Idaho, Nevada, and New Mexico. Following legislative changes to California's community property law, the Court allowed Californians to begin splitting community income starting in 1931.⁵⁷ Thus beginning in 1931 there was a sharp disparity in the tax treatment of married couples' incomes between the eight community property states and common law states. This led to a married couple in community property states usually having lower taxes than the same two persons if they were not married, a difference that largely did not exist in the common law states.⁵⁸

The disparity between community property and common law states persisted until Congress enacted joint filing in 1948.⁵⁹ As enacted, joint filing extended community property-style income splitting to all of the common law states. There was considerable activity preceding the 1948 Act, however. Some in Congress had tried to reinstitute uniform income tax treatment

⁵⁵ In *Robbins*, Justice Sutherland dissented (without opinion) and Justice Stone recused himself. Between the cases, Chief Justice Taft resigned, and Justice Roberts, who wrote the unanimous opinion in *Seaborn*, was appointed. In *Seaborn*, Chief Justice Hughes and Justice Stone (again) recused themselves.

⁵⁶ See Bittker, *supra* note 3, at 1408.

⁵⁷ See *United States v. Malcolm*, 282 U.S. 792 (1931).

⁵⁸ As discussed in note 46, La Lumia's work suggests that the vast majority of higher income couples in common law states did not shift significant capital income to wives. However, many of the very wealthiest couples did so, and the Treasury at the time believed inter-spousal income shifting accounted for a quite substantial loss of tax revenue. See Ventry, *supra* note 33, at 1511 (citing for example *Tax Evasion and Avoidance: Hearings Before the Joint Committee on Tax Evasion and Avoidance*, 75th Cong. 310 (1937)).

⁵⁹ For a history of the 1948 Revenue Act and its impact on the taxation of the family see Stanley S. Surrey, *Federal Taxation of the Family: The Revenue Act of 1948*, 61 HARV. L. REV. 1097 (1948). Surrey called the enactment of joint filing the "one bright spot" in the Act. *Id.* at 1106. Technically we should say that the 1948 Act allowed for the first practical joint filing. Joint filing had been allowed since the early days of the income tax, but because the tax-rates for joint filers were the same as for individuals this made joint filing very unappealing since progressive rates ensured that a couple's taxes would almost always be higher with a joint return.

for community property and common law states on a number of occasions, as early as 1920 and most notably during 1941,⁶⁰ but all attempts failed.⁶¹ In 1939, Oklahoma enacted an optional community property system, attempting to get the same tax treatment as its neighbor Texas, to which Oklahoma claimed it was losing wealthy taxpayers.⁶² The Treasury refused to recognize Oklahoma's law and was upheld by the Supreme Court in 1944 because Oklahoma's regime was optional.⁶³ Not to be deterred, Oklahoma enacted another community property regime in 1945, this one mandatory. The Treasury thus allowed married Oklahomans to begin splitting their income in 1946. It likewise recognized a similar law in Oregon in 1947. When Congress enacted the 1948 Revenue Act, Michigan, Nebraska and Hawaii had very recently passed community property statutes that were as yet unrecognized by the Treasury.⁶⁴ Indeed, joint taxation was passed in 1948 in large part to prevent the remainder of the states from simply opting into community property "impetuous[ly]" to get the tax benefits.⁶⁵

It is important to note that the income tax changed substantially between the *Seaborn* decision in 1930 and the 1948 Revenue Act. In 1930 the income tax was paid only by the rich,

⁶⁰ See Bittker, *supra* note 3, at 1409. The proposed bill in 1941 would have enacted mandatory joint filing, but with rates for joint income the same as for single taxpayers. This form of mandatory joint filing would have created uniform taxation across community property and common law states, but at the same time imposed a very substantial marriage penalty for two earner couples.

⁶¹ Congress found it very difficult to pass any legislation removing the favorable treatment from the community property states. In hearings on the 1948 bill Senator Edwin Johnson stated in hearings on the Act that "[y]ou cannot take it away from the [community property states]; we have tried that." (quoted in Surrey, *supra* note 59, at 1105). Likewise, Senator Eugene Milliken of Colorado stated to an expert testifying before the Senate who proposed overturning *Seaborn* legislatively: "The difficulty is that it is not a novel thought. It has been tossed in the hopper around here a number of times. But legislatively it has not been possible to do." See *id.*

⁶² Stephanie Hunter McMahon, *To Save State Residents: States' Use of Community Property for Federal Tax Reduction*, 27 L. & HIST. REV. 585, 593 (2009).

⁶³ *Commissioner v. Harmon*, 323 U.S. 44 (1944).

⁶⁴ Pennsylvania also passed a community property statute in 1947, but the state supreme court ruled that the statute violated the state constitution. See Note, *Epilogue to the Community Property Scramble: Problems of Repeal*, 50 COLUM. L. R. 332, 334 (1950). The five states and Hawaii that had switched to community property quickly reverted after the 1948 Revenue Act. Carolyn Jones argues that concerns that community property gave women too many rights—undermining traditional gender roles—spurred the return to separate property. Carolyn C. Jones, *Split Income and Separate Spheres: Tax Law and Gender Roles in the 1940s*, 6 L. & HIST. REV. 259 (1988).

⁶⁵ S. Rep. No. 1013, 80th Cong. 2d Sess. 25 (1948).

as it had been since the tax's inception in 1913. About 2.4 million relatively wealthy Americans filed returns in 1930.⁶⁶ The nation's insatiable need for revenue to fight World War II, however, turned the income tax into a mass tax. In 1939 the government collected about \$1 billion through the individual income tax, in 1945 it raised more than \$19 billion, an increase of more than 14 fold after adjusting for inflation.⁶⁷ The government achieved this increase in revenue by both increasing the number of people paying the tax (through lower exemptions) and by substantially increasing rates. By 1947, nearly 55 million Americans were filing returns. In 1930 the Code imposed a 1.5% tax on the first \$4,000 of ordinary taxable income after exemptions, and 3% on the next \$4,000. The comparable rates in 1948 were 23% and 33% respectively.⁶⁸

C. Size of Income-Splitting Tax Advantage

The Treasury rulings (1920-1926) and *Seaborn* (1930) made marriage tax-advantaged for higher income couples in community property states, but this advantage was small for all but the richest citizens until the run up to World War II. For example, even for a reasonably well off single-earner couple with today's equivalent of \$100,000 in income, the marriage bonus was moderate or close to zero (depending on the assumptions used concerning the couple's pre-marriage taxes) during the period from *Seaborn* until 1940.⁶⁹ This bonus increased

⁶⁶ See Annual Report of the Commissioner for Internal Revenue for 1930, 1931, 1932 (Tables displaying number of returns processed from each tax year).

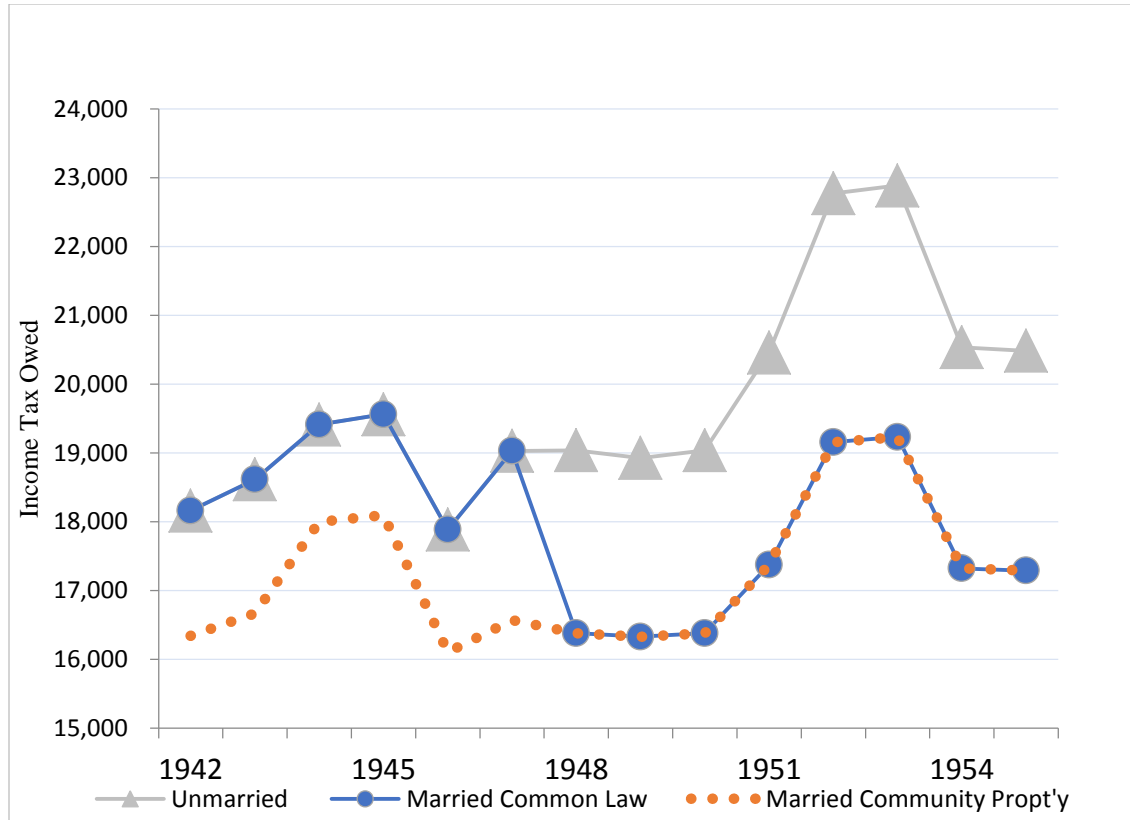
⁶⁷ Annual Report for the Commissioner for Internal Revenue 1946, 1 ("individual income tax and withholding), both figures are nominal. The cumulative inflation rate between 1939 and 1946 was about 29%.

⁶⁸ Based on 1948 rates and using the CPI to update for inflation <http://data.bls.gov/>

⁶⁹ Assuming the husband could claim his future wife as a dependent before marriage, there would be basically no marriage bonus at all because with two personal exemptions, the husband would be taxed in the first bracket only. If the spouse's personal exemption was previously unused, the bonus from *Seaborn* until 1940 averaged \$1,100.

substantially, however, starting in 1941, and averaged \$3,000 per year from 1948-1953 (or as high as \$5,000 on other assumptions).⁷⁰ This is shown in Figure 2:

Figure 2: Taxes Owed by a Single-Earner Couple with \$100,000 in Income



In 1947, the Treasury estimated that extending income splitting to common law states would cost \$743.5 million (or \$8.2 billion in 2016 dollars).⁷¹ This represented a 5% decline in the *total* individual income taxes collected from common law states.⁷² Treasury's prediction appears to have been roughly accurate: after the 1948 Act, revenues from common law states

⁷⁰ Figure 2 assumes the future wife used her full personal exemption prior to marriage. If not, then the average bonus from 1948-1953 was just over \$5,000. For simplicity, I do not adjust Figure 2 for small marriage bonuses built into the structure of the personal exemptions in the years 1942-1944, discussed *supra* in note 39, and I also ignore the question of whether the value of the wife's personal exemption—which could be used by the husband after marriage—would have been greater after the couple wedded.

⁷¹ John D. Morris, *Inequities in Tax Cited by Treasury*, N.Y. TIMES, June 19, 1947 at 4.

⁷² See Annual Report of the Commissioner of Internal Revenue for Internal Revenue for Fiscal Year Ended June 1947.

declined 4 percentage points more than in community property states.⁷³ Full income splitting for married couples remained in place until 1969, when Congress moved toward the current system. The 1969 changes reduced marriage bonuses for single-earner couples, and introduced the first marriage penalties for dual-earner couples with similar incomes.⁷⁴

The differences for married couples between individual taxation and joint taxation with income splitting were not just economically important; they also had salience in the popular press. The *Wall Street Journal* featured an article on the results of *Seaborn*.⁷⁵ The *Los Angeles Times* featured a breathless front page story about *United States v. Malcolm*, which allowed married California couples to split incomes in 1931.⁷⁶ Similarly, Oklahoma and other states' attempts to obtain income splitting for their residents by switching to community property focused attention on the issue, garnering articles in articles in *Business Week*, *Colliers*, *Harper's*, *Newsweek*, and *Time* among others.⁷⁷ Not surprisingly, the 1948 extension of income splitting to all common law states garnered even more coverage because it affected a much greater portion of the population and the Act created much larger savings for those couples than did *Seaborn*.

⁷³ See Annual Report of the Commissioner of Internal Revenue for Internal Revenue for Fiscal Year Ended June 1949. This figure is derived from comparing revenues collected from June 1948 to June 1949 relative to June 1947 to June 1948—the January to June portion of 1948 should have been largely unaffected the Revenue Act of 1948 since 1948 taxes would not be collected until 1949. Even withholding would not have been affected much since the Revenue Act was not passed until April 2, 1948.

⁷⁴ See *Druker v. Commissioner*, 697 F.2d 46, 48 (2d Cir. 1982) (detailing a history of the legal challenges to the “singles penalty/marriage bonuses”).

⁷⁵ *Taxes on Community Income*, WALL ST. J., Dec. 22, 1930 at 2; see also *Income Tax Facts*, WASH. POST, Feb. 4, 1931 at 11.

⁷⁶ *Community Tax Ruling Studied: Married Man's Salary Held Property of Both*, L.A. TIMES, Jan. 21, 1931 at A1 (reporting the holding in *Malcolm* even before the full text of the decision had arrived on the West Coast).

⁷⁷ See articles in *Business Week*, *Colliers*, *Harper's*, *Newsweek*, and *Time* among others cited in Carolyn C. Jones, *Split Income and Separate Spheres: Tax Law and Gender Roles in the 1940s*, 6 L. & HIST. REV. 259, 268 n.67 (1988). Stanley Surrey, writing in 1948, also noted that “[t]he preceding [history concerning income-splitting is] familiar even to non-tax experts, for this phase of federal taxation has been widely discussed.” See Surrey, *supra* note 59, at 1104.

During the first tax season when residents of common law states could split incomes, numerous papers featured stories about the tax savings allowed by joint filing.⁷⁸

DATA

The data used in this study are drawn from the 1960 decennial census, available as part of the Integrated Public Use Microdata Series (IPUMS). The 1960 census asked 25% of the population about when they married for the first time. The publicly available sample is built solely from this portion of the population, which allows me to create a year-by-year record of first marriages from 1942 to 1953 for about 1 million people.^{79,80} Obtaining yearly data is important because it allows me to look at changes in marriage rates over this relatively short window around the 1948 Revenue Act, making my estimate of the impact of the tax change more reliable than looking at changes between one decennial census and another.⁸¹

Below, I run my analysis using each of three measures to proxy for those with enough income to be affected by the 1948 Act: education, 1960 income projected back to income around 1948 by adjusting for inflation and economic growth, and a prediction of income in the period around 1948 based on a person's occupation, education, age, and state of residence. All the

⁷⁸ See, e.g., *Permit Marital Tax Split In All States In Filing: Results in Lower Levy in Most Cases*, CHICAGO TRIBUNE, Feb. 3, 1949 at C7; *Tax Angles: Joint Returns Save Money for Many*, BOSTON GLOBE, Feb. 4, 1949 at 18; *Suggestions On Income Tax Returns: Joint Report Works Out Nearly Always Cheapest Procedure*, Hartford Courant, Feb. 22, 1949 at 3; *Joint Return Eases Family Income Taxes: Offers Privilege \$3,600 Starting Point One Return for Two*, CHRISTIAN SCIENCE MONITOR, Dec 7, 1948 at 2. Interestingly, the instructions to the individual tax form from 1930 to 1947 do not appear to have discussed community property income splitting. Perhaps this was Treasury's small way of thumbing its nose at *Seaborn*.

⁷⁹ The 1950 census only requested marriage dates from the 5% of citizens who received the long-form. Since the publicly available portion is only 1% of this 5%, there is not enough data to construct year-by-year measures for a large enough population. Moreover, the effect of the 1948 Revenue Act is likely to have extended beyond 1950.

⁸⁰ Some individuals died between 1948 and 1960. However, this will primarily be a problem if there is differential mortality between common law and community property states and this differential mortality is correlated with changes in the marriage rate across income groups. I see no reason to think this is likely, especially because those most likely to be affected by the policy were in their 20s or early 30s around 1948 and therefore were quite likely to survive until 1960.

⁸¹ I can only observe year-by-year changes when people entered their first marriage, rather than all marriages.

measures produce fairly comparable results, but in my main specification I use education levels. This is in keeping with LaLumia's related work, which uses college education as an indicator for those affected by income splitting.⁸² Married men with at least two years of college education on average earned about \$5,000 in 1950, enough to generate important tax savings from income splitting with a spouse who did not work or who was paid significantly less.⁸³ By contrast, the median education for married men in 1950 was having completed 9th grade. Men who completed 9th grade or less, earned \$2,500 per year on average with a substantial majority not making enough to benefit from income splitting. I use these education levels as the primary definition of high and low income.⁸⁴

I focus on the marriage rate of men because I can use various measures of income to confirm my primary results. That is not possible for women because the majority of women in higher income households were not in the labor force. The marriage behavior of women, however, closely tracks that of men. I examine marriage rates among women in the robustness section, using high-school graduates as a proxy for those women who are most likely to enter into a marriage where income splitting is valuable.⁸⁵

⁸² See LaLumia, *supra* note 31.

⁸³ 1950 Public Use Census Data (author's calculations).

⁸⁴ This means that in the primary specification "middle" education men who completed 10th grade but less than two years of college are excluded as too likely to contain both a significant fraction of men who are affected by income splitting, and a large proportion who are not. Regardless, the results are not sensitive to other ways of defining high and low income to include those men in the analysis. As shown in the robustness section, the results are similar if high school graduates are considered higher income, and all those who did not complete high school are deemed lower income.

My measures of who was affected by the Act are imperfect, and will include some individuals who did not have enough income to be affected. This will "attenuate" my results toward 0. Put differently, this means that my results should be a lower bound on the actual effect of the Act.

⁸⁵ In 1950 more than half of married female high-school graduates lived in a household which would benefit from income splitting, while far fewer of those who did not graduate from high school lived in such a household.

A. Summary Statistics and Graphs of Differences

Table 1 presents summary data on men in common law and community property states immediately before the 1948 Revenue Act. Overall residents of the two kinds of states look quite similar. It is true that men in community property states were somewhat better educated, having completed a bit less than half-a year of additional schooling on average. One should not be overly concerned that the education figures are not identical, however. To start with, the high income group looks very similar across the two types of states. Moreover, differences between the residents of the two types of states will not bias the estimates of the model run below if one fully controls for them in the regression. Indeed, even assuming that fully controlling for the differences between the residents of the two types of states in the regression is not possible, the model will still produce unbiased results so long as those differences are largely consistent through the twelve-year period of the study. Economists refer to this as a “common trend” requirement. There is good reason to believe in a common trend here: in the six years prior to the Act there was a 90% correlation between the marriage rates among high income men across the two types of states, and a 92% correlation⁸⁶ after adjusting for the marriage behavior of low income men.

⁸⁶ = $\text{corr} \left((MR_{\text{common,hi inc}} - MR_{\text{common,low inc}}), (MR_{\text{community,hi inc}} - MR_{\text{community,low inc}}) \right)$ from 1942 to 1947 where $MR_{i,j}$ represents the marriage rate in state type i (common or community) of men of income type j (high or low income).

**Table 1:
Summary Statistics
Adult Men in 1947**

	Community Prop.	Common Law
All Adult Men		
% White	90.6%	90.7%
% Aged 16-19	5.7%	5.2%
% Aged 20-24	14.3%	13.5%
% Aged 25-29	14.6%	13.7%
% Aged 30-34	13.2%	13.0%
% Aged 35-39	12.1%	12.0%
% Aged 40-49	19.6%	20.0%
% Aged 50-59	12.5%	14.1%
% Aged 60+	8.0%	8.6%
% Ever Married	73.7%	73.2%
Avg Years Education ¹	9.94	9.56
% 2 years of College +	18%	14%
% 10th Grade - 1 Yr College	38%	35%
% 9th Grade or less	44%	51%
Number of Men	79,452	342,435
High Income Men²		
% White	95.8%	96.3%
% Aged 16-19	8.4%	8.1%
% Aged 20-24	20.9%	20.0%
% Aged 25-29	18.7%	17.9%
% Aged 30-34	14.2%	13.4%
% Aged 35-39	11.3%	11.3%
% Aged 40-49	14.8%	16.9%
% Aged 50-59	7.7%	8.4%
% Aged 60+	4.0%	4.0%
% Ever Married	62.5%	61.5%
Number of Men	14,013	47,079

Source: 1960 census iPUMs sample. Adults defined as persons 16 or older.

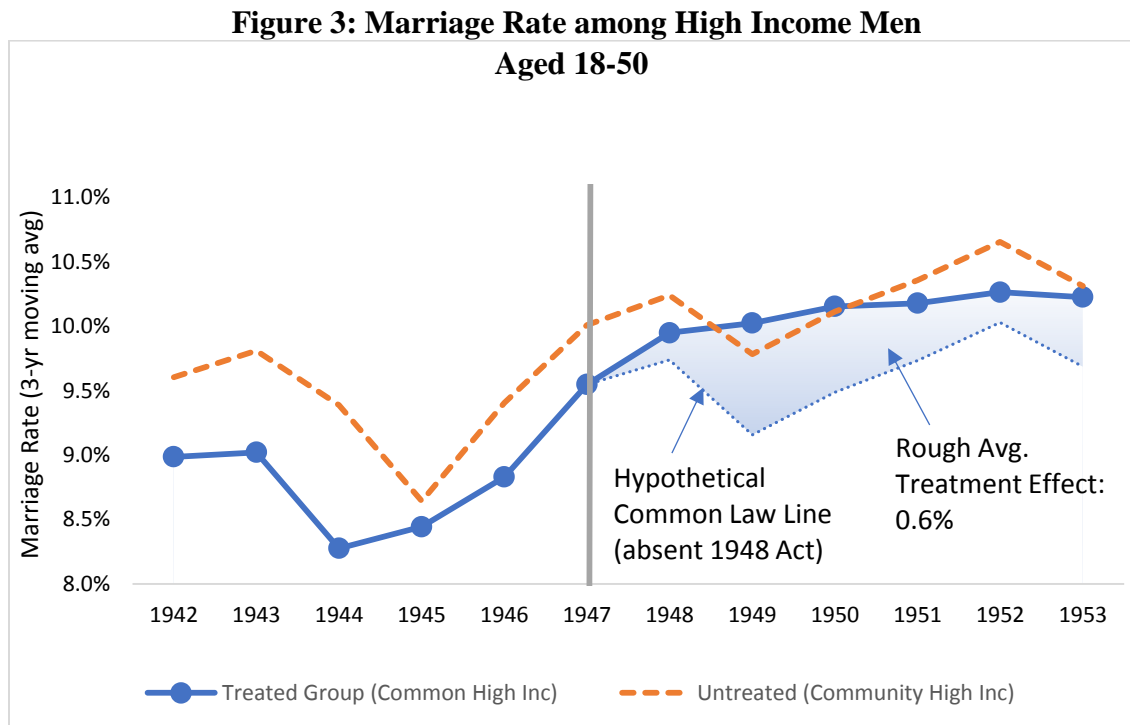
¹ Counting 1st grade as the first year of education.

² Defined as men who completed at least 2 years of college.

It is easier to verify or reject a common trend visually. Figure 3 plots the marriage rate among high income men aged 18-50.⁸⁷ The dashed orange line is the marriage rate in the

⁸⁷ I take a running three year average to smooth the rates and make patterns easier to see by reducing noise.

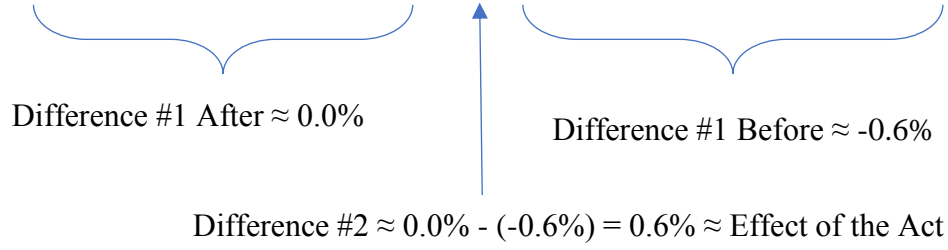
untreated, community property states. The solid blue line (with circular markers) is the marriage rate in the treated, common law states.



The rates evolve in roughly parallel fashion in the years before 1948, but the common law line sits about 0.6% below the community property line. The small blue dotted “hypothetical common law line” represents roughly what one would expect in the common law states absent the 1948 Act: to continue to be about 0.6% below the community property line. Instead, following the Act, the common law line quickly rises to about even with the dashed line in the period after the Act (1948-1953). The shaded blue area between the actual common law line and the hypothetical represents an approximate estimate of the effect of the Act, about 0.6%.

Figure 3 is a graphical depiction of a “difference in difference” model because it allows us to visually take the difference (subtract) the difference in marriage rates of high income men across the types of states before the Act, from that after the Act:

$$Effect \approx (MR_{common, Post48} - MR_{community, Post48}) - (MR_{common, Pre48} - MR_{community, Pre48})$$



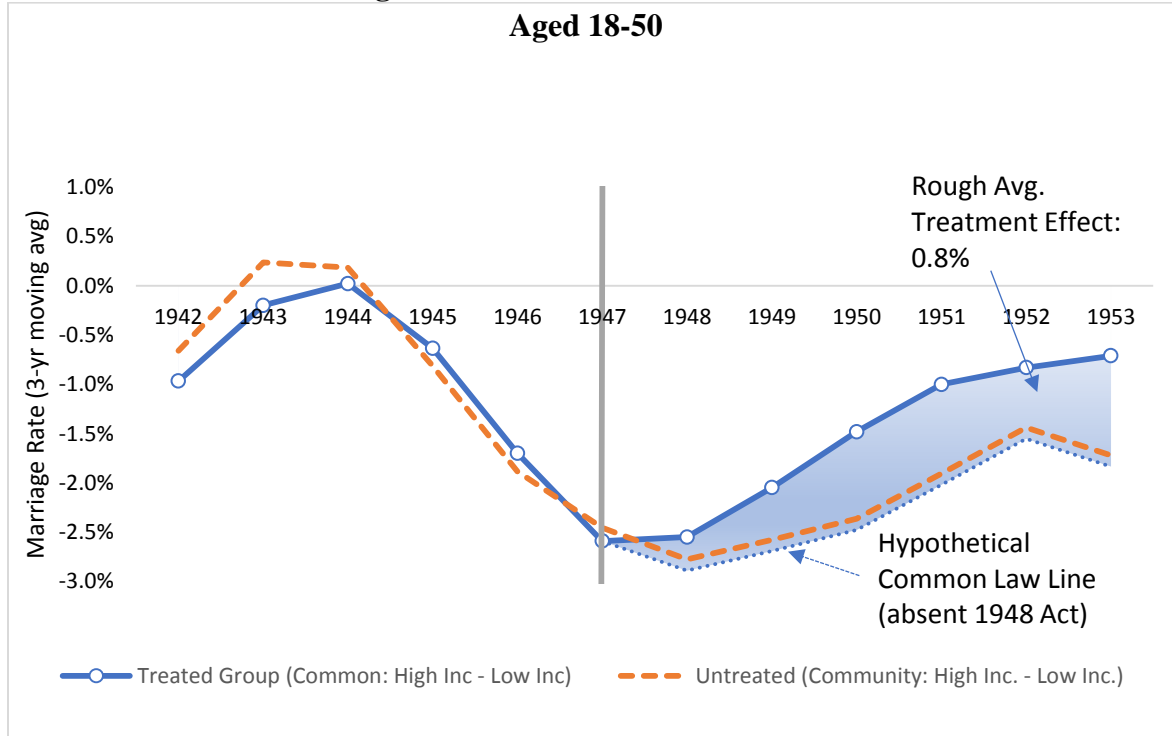
where $MR_{common, Post48}$ stands for the marriage rate among high income men in common law states after the enactment of the 1948 Act, etc.

Although the common trend appears to roughly hold in Figure 3, one can use the data on marriage rates among lower income men to help control for any remaining confounding factors. Figure 4 therefore plots the marriage rate across the two types of states of *high income men relative to low income men*. Thus, the orange dotted line represents the community property states and is equal to

$$(\text{MarriageRate}_{community, high inc men} - \text{MarriageRate}_{community, low inc men}).$$

The solid blue line (with the circular markers) represents the comparable figure for common law states.

**Figure 4: Marriage Rate of
High Income Men - Low Income Men
Aged 18-50**



In Figure 4, the common law and community property lines track each other very closely in the pre-1948 period, with the common law line on average 0.1% below the community property line. They separate, however, starting in 1948 with the common law line rising faster from 1948-1953. Again, the small dotted blue line after 1948 represents approximately what one might have expected in common law states absent the Act: to continue to track the community property line almost exactly. Instead, the actual common law line rises after the Act, sitting on average 0.7% above the community property line. The shaded blue area again represents an estimate of the effect of the Act—about a 0.8% increase in the marriage rate among treated men (after controlling for the untreated states and behavior of lower income men).

Figure 4 represents a visual “difference in difference in difference” (or triple difference) model. In Figure 3 there were two “differences” in the marriage rate of high income men: (1) before and after the Act, and (2) common law and community property states. Figure 4 adds a

third difference in marriage rates: (3) of high and low income men. Intuitively this can still be thought of as a difference in difference model as in Figure 3, except that instead of using the raw marriage rate among high income men, the triple difference model uses an adjusted measure of the marriage rate, which controls for confounding factors using the behavior of low income men.

Both Figures 3 and 4, while useful, use aggregated data. The formal model below implements the triple difference model implicit in Figure 4, but uses individual level data to control for a variety of other factors which might affect marriage rate.⁸⁸

FORMAL MODEL AND RESULTS

I estimate a model which measures how quickly an unmarried man will marry for the first time after he turns 18.⁸⁹ For an unmarried individual i , living in state j , in age-group k ,⁹⁰ in income group l ,⁹¹ in year t (1942-1953), the probability of marrying is:

$$\Pr(\text{Marr}_{i,j,k,l,t}) = \gamma_{j,k,l} + \kappa_t * I_{[t=\text{year}]} + \eta * I_{[\text{Post}'48 * \text{Hi Inc}=1]} + \lambda * I_{[\text{Post}'48 * \text{Common}=1]} + \beta * I_{[\text{Post}'48 * \text{Hi Inc} * \text{Common}=1]} + \pi * X_{\Delta \text{ State-j Per Capita Inc}} + \rho * X_{\Delta \text{ State-j Emp}} + \mu_i$$

I use state-age-income group fixed effects (γ). In practice, this restricts my analysis to *changes* in marriage rates following the 1948 Act within small groups that should be highly comparable (e.g., 20-24 year-old high income men in Michigan). Using these fixed effects reduces the likelihood of the results being driven by factors unrelated to the 1948 Act. To see this, assume, for example, that high income men in Michigan, marry especially early throughout the period of study, compared to those in other states. Let us also assume that for some

⁸⁸ Formally running the difference in difference model implicit in Figure 3 using individual level data yields comparable results to the ones presented below.

⁸⁹ This model is closely related to what economists call a “proportional hazard” model, which is used to estimate the probability an event will occur during a given time period, where that event can occur at most one time for a given individual (e.g., death of an individual, or dissolution of a firm etc.).

⁹⁰ The age-groups are 16-19, 20-24, 25-29, 30-34, 35-39, 40-49, 50-59, 60+

⁹¹ Divided into higher income (=1) (at least two years of college) and lower income (=0) (9th grade or less) as described in Table 1.

demographic reason—i.e., a reason totally unrelated to the passage of the Act—there was a spike starting in 1948 in the number of high income young men in Michigan. Without the fixed effects, this demographic spike could bias the results: the spike in young high income Michiganders (who tend to marry especially early) would lead to an increase in the marriage rate after the Act among high income men in common law states as a whole and thus to the estimated treatment effect. Using fixed effects, however, the model looks only at changes in marriage rates *within* state-age-income groups, so the spike in young high income Michiganders would have no effect on my results. This is because there has been no change in the marriage behavior *within* any state-age-income group, (just a change in the number of young high income Michiganders, who married especially early both before and after the Act).

In addition, $\pi_{\Delta \text{ State-}j \text{ Per Capita Inc}}$ and $\rho_{\Delta \text{ State-}j \text{ Emp't}}$ control for yearly changes in state per-capita GDP and unemployment. This helps to control for economic changes which can affect marriage rates.⁹²

I also control for the national trend for each year (κ_t). This soaks up the effect of any factor that uniformly affects the marriage rate of all men across the nation.⁹³ I likewise control for any national changes that uniformly influence high income men relative to low income men following the 1948 Revenue Act (η). Similarly, I control for factors that affect uniformly all

⁹² In fact, the inclusion or exclusion of this set of controls ends up making relatively little difference, presumably because these factors largely affect high and low income men similarly and hence are controlled for by inclusion of low income men as well in the model.

⁹³ Note that the national trend in marriage rates in a given year (κ_t) is defined as having a “uniform” effect on all men somewhat tautologically. A factor—say a change in the average age at which men enter the work force—need not actually affect all men *identically* to contribute to the national trend κ_t . Instead (roughly) the average nationwide effect of the change in age of workforce entry on marriage rates will be reflected in the national trend. The remaining effect of this change in age of workforce entry on different men will appear elsewhere in the model. The same is true of the “uniform” effect of factors controlled for in η (factors uniformly affecting high income men relative to low income men after the Act) and λ (factors uniformly affecting common law men relative to community property men after the Act).

men in common law states differently than those in community property states following the Act (λ).

Those three factors (κ_t , η , λ) are implicitly already controlled for in Figure 4. Recall that in Figure 4 the lines for common law and community property states represent the *relative* marriage rates among higher and lower income individuals. Because the national trend κ_t is defined as affecting *all men uniformly* (including both high and lower income men), changes in the national trend will leave relative marriage rates unchanged in both types of states. Similarly, consider changes in factors which uniformly affect common law men after the 1948 Act (λ). Because these factors, by definition, affect high and low income men in common law states the same way, the relative marriage rate will be unaffected. Last, consider factors that affect high income men identically across both types of states after the Act's passage (η). These factors will change the relative marriage rates in both types of states because they affect high income men differently than low income men. These factors, however, will change relative marriage rates *in exactly the same way in both types of states*. Thus the difference between the two lines will stay the same.

Finally, we come to the coefficient of interest $\beta_{\text{Post}'48*\text{High Inc}*Common}$, which estimates the effect of the 1948 Revenue Act on marriage rates. $\beta_{\text{Post}'48*\text{High Inc}*Common}$ formally measures the treatment effect shaded in Figure 4.

Table 2:
Estimated Impact of 1948 Revenue Act on Marriage Rate: Men 18-50

Post 1948*High Income	-0.74% (0.27)***
Post 1948*Common Law State	-0.12% (0.37)
Post 1948*Common Law State*High Income	0.82% (0.31)**
R^2	0.03
N	695,348

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Estimated year, state-age-income fixed effects, and state level economic controls not reported, standard errors clustered at the state level

This .82 percentage point increase in the likelihood a high income, never-married man in a common law state marries for the first time in a year corresponds to those individuals marrying on average about 5 months sooner. In all the robustness checks this figure is at least 3 months and every specification implies that tens or hundreds of thousands of Americans who were married at any given time during this period, who would not have been married if not for joint taxation with income splitting.

A. Robustness of the Main Results

Table 3:
Impact of 1948 Tax Change on Marriage Rates

	WWII Control	South & West Regions Only	High Income = HS Grad, Low Inc = Not	High and Low Inc based on 1960 Income ¹	High and Low Inc based on Predicted Income ²	Using Adjusted Location ³	Women 18-50 Hi Inc = HS grad, Low Inc= Not
Post'48*High Income	-0.70% (0.26)**	-0.74% (0.27)**	-0.35% (0.19)*	1.52% (0.18)***	1.99% (0.20)***	-0.42% (0.22)*	-0.45% (0.32)
Post'48*Common Law	-0.17% (0.33)	-0.27% (0.39)	0.04% (0.39)	-0.03% (0.24)	0.21% (0.43)	0.09% (0.33)	-0.13% (0.29)
Post'48*Common Law*High Income	0.79% (0.31)**	1.05% (0.43)**	0.51% (0.22)**	0.69% (0.27)**	0.79% (0.48)*	0.52% (0.29)*	0.74% (0.36)**
High Mobilization*Post'48	-0.21% (0.18)						
R ² (excluding fixed effects)	0.03	0.03	0.03	0.04	0.03	0.03	0.04
N	695,348	338,696	1,252,964	927,873	886,726	552,414	966,312

Estimated year, state-age-income fixed effects, and state economic controls not reported, standard errors clustered at the state level.

¹ 1960 income is converted to income during the study period by deflating it and accounting for per-capita economic growth. High income is defined as making at least \$4,200 per year and low income less than \$3000 in 1948 dollars.

² A person's predicted income is based on a regression of 1960 income on a variety of fixed effects: one for each level of education, each state, each census defined occupation, and each age (in years), for unmarried individuals. This gives a predicted income for each person in each year under study. High and low income are defined as described in note 1.

³ The adjusted location is based on the following hierarchical rule: (1) if a man lives in the state in which he was born, then that state is used; (2) if a man marries and he and his spouse reside in his spouse's birth state, that state is used; (3) if a man marries and both he and his spouse are from the same birth state, then the birth state is used; (4) if the man migrated across state lines in the last 5 years, his birth state is used. Men whose location is placed abroad are not included in the analysis.

Table 3 shows the results of various robustness checks. The first column deals with some of the potential problems raised by the fact that some states sent a higher percentage of men off to World War II than others.⁹⁴ We would expect marriage rates to rise faster in states with higher mobilization once the war was over.⁹⁵ Nevertheless, including this control does not significantly change the estimate of the impact of the 1948 Revenue Act. This makes sense because state specific factors, like mobilization, are already controlled for in the main estimate as long as the factors affect high and low income individuals similarly. The closeness of the estimates with and without the World War II control also suggests that the GI Bill is unlikely to be driving the primary result. The second column shows that although the community property states are grouped in the South and West, the results are not driven by forces which are geographically focused and unrelated to the 1948 Act. Instead, when we look only at the difference between the community property and common law states located in the South and West regions of the country, the coefficients are actually slightly larger.⁹⁶ The third, fourth and fifth columns show alternative ways of defining high and low income which all produce roughly comparable results.

The sixth column shows the results of adjusting the location of men in the sample to account for the fact that some may have moved to their 1960 location during or after the study period. Again, this produces fairly similar estimates, as do two alternative ways to adjust for possible migration shown Table A-1 in the Appendix. Finally, the seventh column shows an

⁹⁴ See LaLumia, *supra*, note 31.

⁹⁵ High Mobilization states are defined as having had more than 50% of their men aged 18-44 register for the draft. *See id.*

⁹⁶ In fact, I have also examined just the states of Texas, Oklahoma, Arkansas, Louisiana, Mississippi, Alabama, and Tennessee which should be highly comparable and find similar, actually slightly larger results, albeit less precisely estimated due to the smaller sample size. See Table A-1 in the appendix.

estimate of the treatment effect for women in common law states, where the probability of benefitting from income splitting is based on whether a woman graduated from high school. The effect is similar to that on men in the primary results.

B. Couples Married Younger, Had More Children

Table 4 shows the estimated effect of the 1948 tax change on the age at first marriage among men who marry at some point in the sample period.

Table 4:
Impact of 1948 Tax Change on Age at First Marriage

	Age at first marriage
Post 1948*High Income	0.7047 (0.2017)***
Post 1948*Common Law State	0.3393 (0.1463)**
Post 1948*Common Law State*High Income	-0.4353 (0.2281)*
R^2	0.03
N	67,737

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Estimated year, state-income fixed effects, and control for economic changes and average age state population not reported, Standard Errors clustered at the state level.

As in the primary specification, the coefficient of -.43 years indicates that “treated” men married about 5 months sooner than we would have predicted without the tax change. This effect is large enough to explain almost the entire increase in the probability an unmarried man marries in a given year. Put differently, Table 4 suggests that the tax change induced men who would have married at some point to marry sooner, rather than inducing men who otherwise would not have ever married to marry. This result makes sense because in this era nearly all high income men married eventually: 95% had married by age 40 in 1948. Thus, most of the reaction to the 1948 Act would have to take the form of changes in *when* high income men married, because the pool of high income men who would never have married was quite small.

In addition, higher income men in common law states appear to have had more children following the 1948 Act. The Act could cause couples to have more kids in at least three ways: (1) affected couples married younger thereby increasing the period of marriage during couples' childbearing years, (2) the tax change drove some married women out of the labor force by substantially raising their marginal tax rate, as La Lumia shows, and (3) it increased take-home pay for married couples during the Baby Boom.⁹⁷ The estimated effect of the tax change is that men had .08 more children at home.⁹⁸ On average men affected by the tax change had 2.34 children at home in 1960, so an increase of .08 is about 3.5% increase, a small but important change in the number of children.⁹⁹

⁹⁷ Although the effect of higher income on fertility is theoretically ambiguous, during the Baby Boom it seems likely that increasing income for a given higher income couple would increase the number of children born to that couple.

⁹⁸ Note that I am using number of children in the house in 1960 as a proxy for children ever born. While a substantial fraction of children from marriages early in the period may have left the house by 1960, this should not bias the results unless the difference in the proportion of children who leave the house between higher and lower income households is different between common law and community property states.

⁹⁹ The 1948 Act has theoretically ambiguous implications for divorce among high income couples. On the one hand, the Act provided a new financial incentive not to divorce for treated couples once they married. On the other hand, the Act seems to have caused treated couples to marry younger, which might lead to higher divorce rates because couples had less time prior to marriage to gauge whether they will be good life partners. For contemporary sociological studies finding marriage at a young age was associated with marital instability, see e.g., T. P. Monahan, *Does Age at Marriage Matter in Divorce?*, 32 SOCIAL FORCES 81 (1953); Lee G. Burchinal, *Trends and Prospects for Young Marriages in the United States*, J. MARRIAGE & FAMILY 243 (1965). Note, however, that these studies are not directly applicable because only couples marrying younger than about age 20 were found to be more likely to divorce during this period. The 1948 Act would not have induced many couples this young to marry because few would have had a high enough income to receive a marriage bonus under the Act, and indeed I estimate that the effect of the Act on men younger than 22 is close to 0 and is not statistically significant.

Unfortunately, studying the effect of the Act on divorce empirically is difficult using the existing data. The 1960 census contains some information about whether the respondent had divorced at some point, but no data about when this occurred. This makes it hard to identify the effect of the Act, although it seems very likely the effect was small because this was a period when divorce was still relatively uncommon. (1.3% of men were divorced and not remarried in 1960, along with another 6.9% whose first marriage had ended and who had remarried, but many of these first marriages would have ended because of the death of the spouse rather than divorce). To get a rough estimate of the effect of the Act on divorce, I combine data from the 1940, 1950, and 1960 censuses and compare how the proportion of high income men who were divorced (and not remarried as of the census date) changed over time. In unreported regressions, I find a small, marginally statistically significant uptick following the 1948 Act in the probability a high income man in a treated state was divorced. The data are thus consistent with the Act very modestly increasing the divorce rate, but I emphasize that we cannot make inferences with confidence without more data, as well as a careful look at changes in divorce laws in the various states up to 1960.

Table 5:
Impact of 1948 Tax Change on Fertility

	Number of own children in the household
Post 1948*High Income	-0.0482 (0.0284)*
Post 1948*Common Law State	-0.1128 (0.0590)*
Post 1948*Common Law State*High Income	0.0806 (0.0424)*
R^2 (excluding fixed effects)	0.02
N	67,737

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Estimated year, state income fixed effects, and control for state average age of men not reported, Standard Errors clustered at the state level

DISCUSSION AND POLICY IMPLICATIONS

The main estimate that the 1948 Act caused eligible men to marry three to five months sooner on average is surprisingly large given the social context. Nevertheless, the estimate is plausible in light of the size of the bonuses. A single-earner couple with \$100,000 in income in today's dollars would have received *each year* a marriage bonus about large enough to cover the entire down payment on a house in the new Levittown suburbs around in 1948 (after federal housing guarantees).¹⁰⁰ Or, looked at differently, if that couple's saving rate matched the national average in 1950, the tax benefit of being married would represent about 40% of their annual savings if they did not change their consumption.¹⁰¹ Marriage bonuses of this size allowed couples who were waiting to marry until they could afford a home or to raise kids to

¹⁰⁰ See EDWARD GLAESER, TRIUMPH OF THE CITY: HOW OUR GREATEST INVENTION MAKES US RICHER, SMARTER, GREENER, HEALTHIER, AND HAPPIER Ch. 7 (2011) (noting that a down payment, after FHA guarantees, in Levittown was \$400, although the suburb drew mostly residents with incomes lower than the hypothetical couple).

¹⁰¹ See BUREAU OF LABOR STATISTICS, 100 YEARS OF CONSUMER SPENDING 22 (showing an average savings rate of 10% in 1950), available at <https://www.bls.gov/opub/uscs/1950.pdf>.

marry significantly sooner.¹⁰² Moreover, contemporary observers believed that the tax code was pushing couples to wed as well. Even before the 1948 Act was passed, the Chicago Tribune reported that sociologists predicted joint filing with income splitting would “work for earlier marriages and against long engagements.”¹⁰³ After enactment, newspapers ran articles like “Wedlock Remains Best Tax Relief” and “Bachelors: Don't Forget Dowry Uncle Sam Gives Bride at Altar,” observing that joint taxation induced marriage.¹⁰⁴ Indeed, income splitting crept into popular culture, with a young Jack Lemon proposing to his girlfriend in the 1954 comedy *Phffft* by extoling the virtues of joint filing.¹⁰⁵

As noted above, the main estimates relating to marriage and those regarding children suggest that the role of federal tax policy in increases in marriage and fertility in the Post War period is greater than previously appreciated. The marriage rate among treated men is estimated to have risen 9% (an increase of 0.82% on a base marriage rate of 9.3%) as a result of the Act. Using the main specification, I estimate that at any given time after 1948 there were roughly 200,000 Americans who were married, who would not have been married absent joint taxation with income splitting.^{106,107} These are important effects.

¹⁰² One academic noted at the beginning of the study period: “‘How much money does it take to get married?’ is one of the questions which is most frequently asked by young men and women who are seriously contemplating marriage. What they really want to know is . . . on how small an income and with how little money in the bank they can safely establish a new family. . . . Most intelligent young couples, no matter how anxious they are to start their new families, are seriously concerned, lest they embark on this their most important life venture with economic resources which may prove to be inadequate.” Howard Bigelow, *Money and Marriage* in MARRIAGE AND THE FAMILY (Howard Becker and Rueben Hill eds. 1942).

¹⁰³ *Believe Cupid Will Benefit by Tax Slash Bill*, CHICAGO TRIBUNE, Jan 4. 1948 at 16.

¹⁰⁴ See e.g., Malvina Lindsay, *Wedlock Remains Best Tax Relief*, WASHINGTON POST, July 19, 1954 (“the national trend to wedlock is being given continued reinforcement by the new tax bill [which reserved full income splitting for married couples]”); J. A. Livingston, *Bachelors: Don't Forget Dowry Uncle Sam Gives Bride at Altar*, WASHINGTON POST, Nov 14, 1952; *Investor's Guide: Marriage in December*, CHICAGO TRIBUNE, Dec. 11, 1948 at A6.

¹⁰⁵ Edward Barry, *Movie Romeo Sings of Split in Income Tax: "PHFFFT"*, CHICAGO TRIBUNE, Dec 2, 1954 at C10.

¹⁰⁶ Looked at differently, from 1960 to the election of Ronald Reagan in 1980, a variety of cultural and

The main estimate can also inform our understanding of the effect of marriage bonuses today on middle income single-earner couples. The two changes since the study period which would appear to have the greatest relevance for single-earner couples making \$50,000 to \$110,000 both suggest taxes have a larger effect on marriage decisions of those couples today. Thus, although we cannot know with certainty, it might be reasonable to think of the estimates of the chapter as a floor for the effect of taxes on those couples today. As noted above, the increasing substitutability of marriage and unmarried cohabitation will likely make today's couples more sensitive to tax incentives when deciding whether or when to marry. Around 1948, unmarried cohabitation was quite uncommon among middle and upper income couples. As late as 1965-1974, less than 10% of all women cohabited prior to marriage.¹⁰⁸ Now that figure is over 60%.¹⁰⁹ Moreover, cohabitation has moved slowly from being primarily “a short-term arrangement among childless young adults who . . . quickly break up or marry” to being “accepted [by many] as an alternative to marriage.”¹¹⁰ As cohabitation has become a genuine alternative to marriage for many couples, they will have become more sensitive things like the tax consequences of marriage, which might have been secondary for many of them in the past.

economic factors (including increasing educational attainment and labor force participation by women, loosening attitudes toward pre-marital sex and changing opinions of marriage itself) drove up the median age at first marriage for men, from 22.8 years to 24.7 years, or 1.9 years. *See* CENSUS DEPARTMENT, TABLE MS-2, ESTIMATED MEDIAN AGE AT FIRST MARRIAGE: 1890 TO PRESENT (2017). The estimated effect of joint taxation with income splitting on treated men (driving down marriage age by 5 months) is therefore about 20% the size of the effect of *all the various forces active during the 1960s and 1970s* driving up median age at first marriage for men.

¹⁰⁷ The natural experiment studied here is also promising as what economists term an “instrument” which can help untangle the causal effect of marrying earlier or having more children on health or economic outcomes. I plan to explore this possibility in a follow up paper.

¹⁰⁸ Pamela Smock, *Cohabitation in The United States: An Appraisal of Research Themes, Findings, and Implications*, 26 ANN. REV. SOCIOLOGY 1, 3 (2000)

¹⁰⁹ Wendy D. Manning, *Trends in Cohabitation: Over Twenty Years of Change, 1987-2010*, (presenting data from the National Survey of Family Growth conducted by the CDC).

¹¹⁰ Andrew Cherlin, *The Deinstitutionalization of American Marriage*, 66 J. MARRIAGE & FAMILY 848, 849 (2004). For example, the number of cohabitations which end in marriage within three years dropped from 60% in the 1970s to 33% by the 1990s. *Id.*

Indeed, many other authors have observed that the increase in cohabitation likely amplifies the effect of tax considerations on marriage.¹¹¹

Second, marriage bonuses for single-earner couples making the equivalent today of what the high income couples studied here earned in 1948 (\$50,000 to \$110,000 per year in today's dollars) are bigger now than around 1948. Indeed, the bonuses are at least twice as large now for couples in the \$50,000-\$80,000 range than in the period under study.¹¹² Of course, I should also note that a much lower percentage of couples now have a single earner than in 1948. Still, those couples remain a significant share of the population amounting to tens of millions of people.¹¹³

Americans' responsiveness to tax incentives in making marriage decisions modestly strengthens the case in favor of returning to individual taxation of all married couples.¹¹⁴ One of

¹¹¹ See, e.g., Zelenak, *supra* note 15 at 792-795 (making the rise of cohabitation the central reason that we should be paying more attention to "tax marriage effects"); Chade and Ventura, *supra* note 26, (showing in a formal theoretical model that increasing the acceptability of cohabitation makes the "number of marriages become[] more sensitive to changes in differential tax treatment [of marriage])."

¹¹² The Revenue Act of 1948 may have had particular salience due to widespread coverage of the benefits accruing to community property couples from income splitting and therefore had a larger effect than it otherwise would have. Thus, arguably, the effect of the 1948 Act may not be a floor for today's less salient marriage bonuses, but the rise in cohabitation and increasing size of the bonuses seems likely to outweigh this factor.

It is also worth observing that the effect of taxes on marriage today may be qualitatively different than it was during the study period. Nearly all higher income men in the study period married eventually. Today that is no longer true and the rise in cohabitation may mean that the effect of taxes on marriage will be as much about whether someone ever marries, as when.

¹¹³ Using census data, I calculate that in 1950 about 70% of married couples with total income of \$50,000 to \$110,000 (in 2016 dollars) were single-earner. Based on the 2015 ACS survey, about 35% of married couples with income between \$50,000 and \$110,000 were single-earner. Both figures exclude couples with no one in the work force (e.g., retirees). Note also that while today a lower percentage of married couples are single-earner, a higher percentage of single-earner couples earn enough to be eligible for sizeable marriage bonuses. These effects come close to offsetting in estimating the total proportion of couples affected by large marriage bonuses.

¹¹⁴ This would presumably require Congress to use statutory language that would clearly overrule *Seaborn*. It would also put increased pressure on preventing income shifting within couples. Congress could largely prevent income shifting by adopting a hybrid form of taxation with individual taxation of wages and salary and some form of joint taxation of capital income (e.g., capital income is pooled among spouses and then taxed on a separate schedule or at the marginal rate of the higher earner, etc.). This hybrid system would sacrifice full marriage neutrality, although the non-neutral effects of joint taxation of capital would be concentrated on the wealthy. As explained in note 18, this system would actually follow most of the other countries which have switched to "individual" taxation since 1970.

The problems of inter-spousal income shifting should not be taken lightly. Stephanie Hunter McMahon has argued that the U.K.'s difficulties containing inter-spousal income shifting may well have outweighed the benefits of

the problems with joint taxation is that it distorts individuals' incentives to marry. The greater their responsiveness to these incentives, the greater is the cost in terms of economic efficiency in terms of distorting couple's decisions to marry or not.¹¹⁵ These costs are in addition to the efficiency losses imposed by joint taxation's tendency to raise the marginal tax rate on the lower earning member of the couple, still more often a woman, who is more likely to drop out of the labor force in the face of high taxes than the higher earning member.¹¹⁶ It is of course true that the societal choice between individual and joint taxation of married couples turns on far more than just efficiency. Rather questions of equity are paramount and it is beyond the scope of this Chapter to restate those arguments here. Nevertheless, in the overall societal calculation as to how married couples should be taxed, it is important to learn that joint taxation's distortions on these choices are substantial.

Many people, of course, believe the government *should* encourage couples to marry.¹¹⁷ They might view my primary results as pushing for the U.S. to move back toward the 1948 to 1969 system with joint taxation with income splitting in all brackets. Yet I do not think that

its shift to individual taxation in 1990. See Stephanie Hunter McMahon, *London Calling: Does the UK's Experience with Individual Taxation Clash with the US's Expectations*, 55 ST. LOUIS U.L.J. 159 (2010).

¹¹⁵ See Listokin, *supra* note 3.

¹¹⁶ See e.g., Zelenak, *supra* note 3, at 365-369. As discussed in note 114, observe, however, that joint taxation can also be cast as improving economic efficiency by reducing wasteful tax planning and avoidance in inter-spousal tax transfers and the need to police such transfers.

¹¹⁷ From an efficiency point of view, the government should encourage marriage if there would be positive externalities (or perhaps positive internalities) resulting from the new marriages induced by the policy, which exceed the cost of the program. See Listokin, *supra* note 3 at 195. There is a wide literature showing marriage is associated with greater longevity, family income, and self-reported happiness. See e.g., Robert E. Emery, Erin E. Horn & Christopher R. Beam, *Marriage and Improved Well-Being*, in MARRIAGE AT THE CROSSROADS LAW, POLICY, AND THE BRAVE NEW WORLD OF TWENTY-FIRST-CENTURY FAMILIES (Marsha Garrison & Elizabeth S. Scott eds., 2012). The marriage of a child's parents is also associated with better educational and other outcomes for that child. See e.g., Linda J. Waite, *Does Marriage Matter?*, 32 DEMOGRAPHY 483, 493-495 (1995). The difficulty in this literature is identifying causation. Some of marriage's association with these positive outcomes is likely explained by selection in who gets married and stays married, but it is unclear how much. I am not aware of a study which identifies genuinely exogenous variation in marriage which would untangle this issue. In addition, even if marriage is in fact good for those who have already taken it up, that does not necessarily imply that the same benefits will flow to those who would not have married at that point but were induced to wed by government policies.

conclusion is warranted. Even putting aside the problems joint taxation creates by reinforcing gender roles, etc., the vast majority of the benefit from full income splitting in all income brackets accrues to the very rich. In unreported regressions, I find little evidence that the richest men, who enjoyed the largest marriage bonuses, reacted more strongly than middle class men who saw smaller bonuses.¹¹⁸ This is not terribly surprising: assuming the marginal utility of income declines as men grow richer, wealthy men should react less, per dollar, to changes in tax incentives for marriage.¹¹⁹ Thus for the richest men there are offsetting effects. They are less responsive per dollar of tax incentives, but get the largest marriage bonuses under joint taxation with income splitting in all brackets. This leads to the rich reacting in about the same way as upper-middle income men to joint taxation with income splitting. In addition, dual-income couples get little or no tax benefit from marriage under joint taxation with income splitting. Assuming the government's goal is to encourage marriage in general—given a fixed cost to the fisc—it would be more effective to use individual taxation and provide a fixed subsidy for marriage in the form of a credit to all Americans. This would provide equal incentives for dual income couples to marry and not focus so much of the subsidy on the very rich.¹²⁰

¹¹⁸ Although number of very high income men in the sample is small, making statistical inference less precise, it is possible to rule out at the 95% level that the rich react as much, per dollar of tax incentive, as upper-middle income men.

¹¹⁹ In an economic model of marriage, a person will implicitly trade off the additional utility provided by the tax incentives in favor of marriage, with other factors which would otherwise cause her to marry later absent those tax incentives (e.g., waiting a bit longer to marry to better understand whether she is compatible with her potential spouse). Richer persons will get less additional utility, per dollar, from the tax incentives. By contrast, we would not expect the strength of the factors pushing toward later marriage to vary much by income.

¹²⁰ The results of the paper actually provide some justification for the current structure of marriage bonuses in the income tax for single-earner couples, which provide full income splitting couples making up to around \$150,000 combined, with additional bonuses tapering off thereafter. Nevertheless, if encouraging marriage is the goal, it is unclear why we would use a system under which dual income couples receive little or no benefit (and above \$150,000 face marriage penalties). Moreover, there remain marriage penalties built into many programs designed to help lower income Americans, including those built into the tax code like the EITC.

CONCLUSION

This Chapter has detailed the history of how the U.S. tax system attempted to integrate a tax system based on the individual with community property systems where the family represents the fundamental unit. The inherent tensions in this process created an unusual series of tax changes that provided for joint taxation with full income splitting for married couples in some states well ahead of others. An empirical examination of this natural experiment strongly indicates that tax incentives had an important effect on marriage rates and fertility. This suggests that the effect of taxes on marriage formation and fertility should be accounted for in considering how we should tax families today.

Chapter 3 Economic Crisis and the Integration of Law and Finance: The Impact of Volatility Spikes (with Merritt Fox and Ronald Gilson)

INTRODUCTION¹²¹

Financial economics has become ever more integrated into corporate and securities law, a trend that started decades prior to the financial crisis of 2008 to 2009.¹²² This is as it should be. Corporate and securities law is, at its core, about valuation. Investors provide to a corporation the funds with which it acquires real assets. The investors receive in return financial claims (securities) on the corporation's future cash flows. The size of these future cash flows then depends importantly on management's choice of what real assets to acquire and how well these assets are managed over time. The capital market's pricing of the financial claims acquired by investors is in effect a valuation of these future cash flows.¹²³ Corporate law provides a framework within which a firm's managers make these investment and operating decisions. Properly designed, this legal framework helps spur management to choose and deploy assets in

¹²¹ This chapter was published first in 116 *Columbia Law Review* 325 (2016).

¹²² The interaction of corporate law and finance began in earnest with Michael Jensen and William Meckling's seminal article on the agency costs at the heart of public corporations. Michael C. Jensen & William H. Meckling, *Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure*, 3 *J. Fin. Econ.* 305 (1976); see also Henry Hansmann & Reinier Kraakman, *Agency Problems and Legal Strategies*, in *The Anatomy of Corporate Law* 21 (R. Kraakman et al., eds. 2004) (discussing centrality of agency theory to structure of corporate law). A series of influential articles linking the legal system to corporate finance and the structure of capital markets further evidences how law and finance came to be linked. See, e.g., Rafael La Porta, Florencio Lopez-de-Silanes, Andrei Shliefer & Robert W. Vishny, *Law and Finance*, 106 *J. Pol. Econ.* 1113, 1114–16 (1998) (presenting cross-country study of links between investor legal protection and development of strong capital markets). For an example of current scholarship that is shaped by the intersection of law and finance, see Ronald J. Gilson & Alan Schwartz, *Corporate Control and Credible Commitment*, 43 *Int'l Rev. L. & Econ.* 119, 121 (2015) (analyzing controlling shareholders' extraction of private benefits of control in corporate-finance terms).

¹²³ See Richard A. Brealey, Stewart C. Myers & Franklin Allen, *Principles of Corporate Finance* 80 (11th ed. 2014) (defining market value of debt and equity securities as discounted present value of those securities' future cash flows).

ways that maximize the value of the firm's expected future cash flows. The framework's effectiveness, however, depends in part on the precision with which the capital market prices these financial assets.¹²⁴ Securities law can enhance price accuracy, for example, by mandating that corporations disclose certain information and by regulating the workings of the securities markets and the behavior of those who trade in them.¹²⁵ The better corporate and securities law perform these tasks, the more valuable the corporation's underlying business and correspondingly, the financial claims that the corporation issues.¹²⁶

The 2008 financial crisis raised many questions for financial economics and financial regulation.¹²⁷ Relatively unappreciated so far, however, are the puzzles the crisis poses for understanding how the capital market prices common stocks and in turn, for the intersection between law and finance. During the crisis, there was a dramatic spike, across all industries, in “idiosyncratic risk”¹²⁸—that is, in the volatility of individual firm's share prices after adjustment for movements in the market as a whole.¹²⁹ This phenomenon, it turns out, is not limited to just

124. See Merritt B. Fox, *Civil Liability and Mandatory Disclosure*, 109 *Colum. L. Rev.* 237, 252–60, 264–67 (2009) [hereinafter Fox, *Civil Liability*] (“More information, with the resulting increase in price accuracy, improves the control market's effectiveness in limiting the agency costs of management.”).

125. Consistent with this view, the growing law and finance literature finds that effective corporate and securities law is a precondition for the sophisticated capital markets and corporate ownership structures that appear most closely associated with economic growth. See, e.g., La Porta et al., *supra* note 122, at 1152 (examining evidence and concluding strong investor-protection laws correlate with economic growth). Securities law also can improve pricing accuracy by facilitating the efficiency of capital market microstructure through regulation. We need not address this function here.

126. See Jesse H. Choper, John C. Coffee, Jr. & Ronald J. Gilson, *Cases and Materials on Corporations* 177–78 (8th ed. 2013) (discussing valuation of financial assets).

127. See, e.g., Gary Gorton, *Misunderstanding Financial Crises: Why We Don't See Them Coming* 221 (2012) (reviewing past misunderstandings of financial crises); Robert E. Hall, *The High Sensitivity of Economic Activity to Financial Frictions*, 121 *Econ. J.* 351, 371 (2011) (arguing standard macroeconomic theory does not explain why drops in output and employment persist for significant periods after financial crisis is resolved); Robert E. Hall, *Why Does the Economy Fall to Pieces After a Financial Crisis?*, *J. Econ. Persp.*, Fall 2010, at 3, 3 (2010) (same); see also Andrew W. Lo, *Reading About the Financial Crisis: A Twenty-One-Book Review*, 50 *J. Econ. Literature* 151, 154–55 (2012) (presenting useful reviews of twenty-one books drawn from postcrisis deluge of works analyzing crisis and appropriate responses).

128. See *infra* section I.A (discussing spike in idiosyncratic risk during crisis).

129. See Brealey, Myers & Allen, *supra* note 123, at 174 (defining and comparing “[firm-] specific risk” and “market risk”).

the most recent economic crisis. We show here for the first time that a dramatic spike in idiosyncratic risk has occurred with every major downturn since the 1920s.¹³⁰ This association—between economic crisis and a spike in idiosyncratic risk—is important to the law for two important reasons. First, the spike is caused by a crisis-induced sharp increase in the importance of new information specifically concerning individual firms. Information—its value and its disclosure—is at the center of much of corporate and securities law. Second, idiosyncratic price changes are important because they are at the core of event studies, an econometric technique that, over the last few decades, has moved from an academic tool to assess the impact on stock price of particular corporate actions to providing the central means by which corporate and securities law is applied to specific cases.¹³¹ For example, the Supreme Court recently confirmed the centrality of event studies in securities fraud class actions,¹³² the predominant form of private securities litigation today.¹³³

Now consider the 2008 to 2009 financial crisis’s impact on the stock prices of individual firms. The volatility of individual stock prices increased sharply during the crisis.¹³⁴ A portion of this increase is explained by economy-wide factors like changes in general economic conditions that affect share prices of all stocks. But much more than this was going on. The larger part of the increase in each firm’s overall share price volatility was due to a dramatic rise—five-fold as measured by variance—in idiosyncratic risk, the portion of the volatility that cannot be explained by changes in factors that affect all firms. Rather, there was a large increase, relative to noncrisis

130. See *infra* section I.B (presenting empirical data on crises since 1920s).

131. See *infra* Part III (discussing use of event studies in securities litigation).

132. See *Halliburton Co. v. Erica P. John Fund, Inc.*, 134 S. Ct. 2398, 2415 (2014) (discussing event studies in securities litigation).

133. See Merritt B. Fox, *Securities Class Actions Against Foreign Issuers*, 64 *Stan. L. Rev.* 1173, 1176 (2012) [hereinafter Fox, *Securities Class Actions*] (noting fraud-on-the-market class actions “give rise to the bulk of all the damages paid out in settlements and judgments pursuant to private litigation under the U.S. securities laws”).

134. See *infra* section I.A (presenting empirical data from 2008–2009 financial crisis).

times, in the extent to which an individual firm's share price deviated independently from the change in the market as a whole.

This crisis-related spike in idiosyncratic risk presents a puzzle that is important to understanding how law and finance interact. A firm's share price moves because there is news: a bit of information that changes expectations about the firm's future cash flow and hence the value of the firm's shares.¹³⁵ So for example, the announcement that a defense firm received a profitable government contract that was previously thought possible, but not certain, is news: The chance that the contract would not be received is eliminated. In an efficient market, as soon as a bit of news is revealed, it is promptly and fully reflected in price.¹³⁶ News is thus by definition unpredictable—to be news it cannot be anticipated and thus already incorporated in the firm's stock price—with its importance determined by how much the news changes expectations of a firm's future cash flows. Future price changes are the result of subsequent news.

Some news affects expectations about the future cash flows of most or all firms because it concerns the overall state of economy—for example, the future rate of economic growth, employment levels, interest rates, or inflation. For any given firm's stock, the portion of the total variability in its share price due to such generally applicable news is referred to as the stock's "systematic" risk, or volatility. The remaining variability in the firm's share price is due to news that affects expectations about its particular cash flows and not about the cash flows of most or

135. See Brealey, Myers & Allen, *supra* note 123, at 365 ("[S]tock prices and company values summarize investors' collective assessment of how well a company is doing, both its current performance and its future prospects.").

136. See, e.g., *id.* at 324–25 (noting large body of accumulated empirical evidence that publicly traded issuer shares listed on exchanges such as NYSE and NASDAQ show immediate reactions to revelations of news, after which prices follow random walk).

all other firms in the market—for example news of a labor strike,¹³⁷ an R&D breakthrough,¹³⁸ or a regulatory development.¹³⁹ The literature variously refers to this as “idiosyncratic,” “unsystematic,” “firm-specific,” or “unique” risk, or volatility.¹⁴⁰ Thus, the five-fold increase in idiosyncratic volatility during the 2008 to 2009 financial crisis tells us that each bit of news affecting only a particular firm altered, much more than in normal times, expectations concerning that firm’s future cash flows. In essence, during a crisis, all firms share an increased sensitivity to bits of news that, for each, will not affect most other firms. As a consequence, for example, on any good day for the market as a whole, there are far more big losers than in normal times and on any bad day, far more big winners.

The puzzle is why. An economic crisis concerns problems in the economy as a whole. Why would a crisis suddenly increase the importance of new information that is independent of the effect on the company of news about the overall economy? In this chapter, we take up both the causes of this large, crisis-induced increase in idiosyncratic risk and its implications for how the legal system uses finance, topics yet to be addressed in either the financial-economics or legal literature.

137. See John Dinardo & Kevin F. Hallock, ‘When Unions Mattered’: The Impact of Strikes on Financial Markets, 1925–1937, 55 *Indus. & Lab. Rel. Rev.* 219, 226–32 (2002) (employing event study methodology to find significant negative impacts of labor-strike news on affected firms’ share prices).

138. See, e.g., Marta Falconi & Neil Maclucas, Novartis Boosted by Heart Drug Trials, *Wall St. J.* (Mar. 31, 2014), <http://www.wsj.com/articles/SB10001424052702304157204579473051376848032> (on file with the *Columbia Law Review*) (“Novartis AG said Monday it would seek approval for an experimental heart failure treatment sooner than it had expected The news helped bolster shares of Novartis, which surged 3.31% in late morning trading . . .”).

139. See, e.g., Matthew Curtin, Volkswagen Shares Tumble Following Emissions Allegations, *Wall St. J.* (Sept. 21, 2015), <http://www.wsj.com/articles/volkswagen-shares-driven-lower-1442826436> (on file with the *Columbia Law Review*) (“The auto maker’s nonvoting shares . . . were down 21% in morning trading amid fears of a huge fine in the U.S. as well as long-term damage to the reputation of Europe’s biggest auto maker by sales.”). Stock prices are also affected by background noise in the securities markets—movements that are uncorrelated with any new information. Thus, idiosyncratic risk will at all times include this background noise. For purposes of this Article, we focus on information-based idiosyncratic risk.

140. See Brealey, Myers & Allen, *supra* note 123, at 174 n.25 (“Specific risk may be called *unsystematic risk*, *residual risk*, *unique risk*, or *diversifiable risk*.”).

Part I begins by documenting more fully the link between economic turmoil and idiosyncratic risk. Expanding on earlier work by Campbell et al.,¹⁴¹ we conduct an empirical review, extending back to 1926 and forward to the present. This review shows, for the first time, that *every* major economic downturn in this eighty-five-year period has been accompanied by a substantial spike in idiosyncratic volatility.

Part II seeks to explain why difficult economic times, which are defined in terms of market-wide phenomena, make the future of individual firms more difficult to predict and so make individual stock prices more volatile, independent of the crisis making the overall economy's future performance harder to predict. We canvass several complementary answers. One is that, compared to ordinary times, information about a firm contained in *current* news may become more important in predicting its future cash flows *relative to* the role of the *already existing* stock of knowledge in making such predictions. A second explanation is that the quality of management becomes more important in crisis times. Consequently, the ordinary flow of new information about this subject can cause bigger movements in price because each bit tells the market about something—the quality of management—that the market now regards as more important than it did in ordinary times. A third is that crisis creates uncertainty as to what factors, and hence what information, are important to valuation. Because of this uncertainty, a broader range of information has valuation implications and therefore stock prices move more frequently.

Parts III and IV turn to the implications of our empirical results and of their possible explanations for a number of legal issues that depend analytically on valuation methodology: determining materiality and loss causation in fraud-on-the-market securities litigation,

141. John Y. Campbell, Martin Lettau, Burton G. Malkiel & Yexiao Xu, Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk, 56 J. Fin. 1, 3 (2001) [hereinafter Campbell et al., Empirical Exploration] (finding market-, industry-, and firm-level volatility all “increase substantially in economic downturns and tend to lead recessions”).

determining materiality in cases involving claims of insider trading or of misstatements or omissions in registered public offerings, and determining the extent of deference that should be paid to a corporate board that rejects an acquisition offer at a premium above the pre-offer market price. Litigation with respect to each of these matters has, over the last few decades, increasingly involved the empirical analysis—through the use of event studies—of the idiosyncratic portion of share returns of the companies involved.

The analysis in Part IV yields a number of important insights, starting with ones relating to the use of event studies in fraud-on-the-market class actions, the actions that give rise to the bulk of damages and settlements paid out as the result of private securities litigation. We show that the conventional use of event studies during periods of economic-crisis-induced increased volatility results in understating the number of occasions when a corporate misstatement can be shown to have had a meaningful impact on a firm's stock price. Lowering the standard for statistical significance in crisis times would help to correct for this problem but causes an offsetting problem: With a lower standard, liability is imposed more frequently where the misstatement in fact did not have a sufficient impact on price to justify the costs associated with imposing liability. In the end, this analysis shows that lowering the standard in times of crisis is, on balance, no more likely to improve than to harm social welfare. This suggests that during crisis times, Securities Exchange Commission (SEC) enforcement actions are especially important because they offset the reduced effectiveness of private litigation.

We reach a similar conclusion with respect to insider trading and public offerings of securities. This has particular significance because the information-based explanations that we find most persuasive for why economic crises lead to spikes in idiosyncratic risk suggest that, in crisis times, insiders have substantially more opportunities to profit from trading on the

nonpublic information that they possess and issuers have more opportunities to sell securities at an inflated price.

Finally, “substantive coercion” is rarely a serious justification for target corporation deployment of takeover defenses, but crisis-induced increases in idiosyncratic risk provide an unusually plausible claim that target shareholders may make a mistake in tendering into a hostile offer. However, the example of two cases that raise the issue close in time to financial crises—*Quickturn*¹⁴² and *Airgas*¹⁴³—shows that even a substantive coercion claim based on a crisis-induced spike in idiosyncratic risk is very difficult to demonstrate.¹⁴⁴

I. THE EMPIRICAL RECORD

In this Part, we demonstrate empirically both that there was a sharp increase in idiosyncratic risk during the 2008 to 2009 financial crisis and that this is not an isolated incident. An increase in idiosyncratic risk is associated with poor macroeconomic performance throughout an eighty-five-year period going back to 1926. We show as well that the increase was felt by firms across all industries. Moreover, while some of the increase in idiosyncratic risk is due to the simple fact that the share prices of most firms fell dramatically during the crisis, thereby increasing risk because of the resulting higher firm debt–equity ratios, much of it is not.

142. *Quickturn Design Sys., Inc. v. Shapiro*, 721 A.2d 1281 (Del. 1998).

143. *Air Prods. & Chems., Inc. v. Airgas, Inc.*, 16 A.3d 48 (Del. Ch. 2011).

144. We refer occasionally to an online appendix that contains additional technical information, further empirical results, and demonstrations of the robustness of our findings. See Edward G. Fox, Merritt B. Fox & Ronald J. Gilson, *Economic Crises and the Integration of Law and Finance: The Impact of Volatility Spikes* app. (2015), <https://sites.google.com/site/volatilityspikesappendix/> (on file with the *Columbia Law Review*) [hereinafter Appendix] (providing technical information and further results and demonstrating robustness). For ease of reference, the appendix is also included as Appendix 2 at the end of this dissertation.

A. The 2008 to 2009 Financial Crisis

The 2008 to 2009 financial crisis illustrates starkly the spike in idiosyncratic risk associated with economic crises. From July 1, 2008, to June 30, 2009, firms in the S&P 100 experienced a five-fold increase in the average idiosyncratic volatility, as measured by variance, compared to 2006 to 2007 and a three-fold increase compared to 2007 to 2008. While 2008 to 2009 was a period of enormous general economic turmoil¹⁴⁵—recall that Lehman Brothers failed¹⁴⁶ and AIG was essentially nationalized within seven days in September 2008¹⁴⁷—this increase in idiosyncratic volatility reflects movement in individual firm prices that cannot be explained by the direct impact of general economic conditions. While the most extraordinary increases were among financial firms in the index—forty-fold during the crisis relative to 2006 to 2007—nonfinancial firms increased volatility almost four-fold themselves.¹⁴⁸ Just as dramatically, idiosyncratic risk then returned to approximately normal levels by June 30, 2010. Thus, we see a

145. See Bureau of Econ. Analysis, U.S. Dep't of Commerce, BEA 09-29, Gross Domestic Product, 1st Quarter 2009 (Final), Corporate Profits, 1st Quarter 2009 (Revised) (2009), <http://www.bea.gov/newsreleases/national/gdp/2009/gdp109f.htm> [http://perma.cc/533U-PYGV] (“Real gross domestic product—the output of goods and services produced by labor and property located in the United States—decreased at an annual rate of 5.5 percent in the first quarter of 2009”); Bureau of Econ. Analysis, U.S. Dep't of Commerce, BEA 09-11, Gross Domestic Product: Fourth Quarter 2008 (Final), Corporate Profits, Fourth Quarter 2008 (2009), <http://www.bea.gov/newsreleases/national/gdp/2009/gdp408f.htm> [http://perma.cc/9B73-7NWA] (“Real gross domestic product . . . decreased at an annual rate of 6.3 percent in the fourth quarter of 2008”); David M. Herszenhorn, Bailout Plan Wins Approval; Democrats Vow Tighter Rules, N.Y. Times (Oct. 4, 2008), <http://nytimes.com/2008/10/04/business/economy/04bailout.html?pagewanted=all> (on file with the *Columbia Law Review*) (noting former Treasury Secretary Henry Paulson in 2008 urged “emergency bailout legislation with a warning that the American economy was at risk of the worst economic collapse since the Depression”).

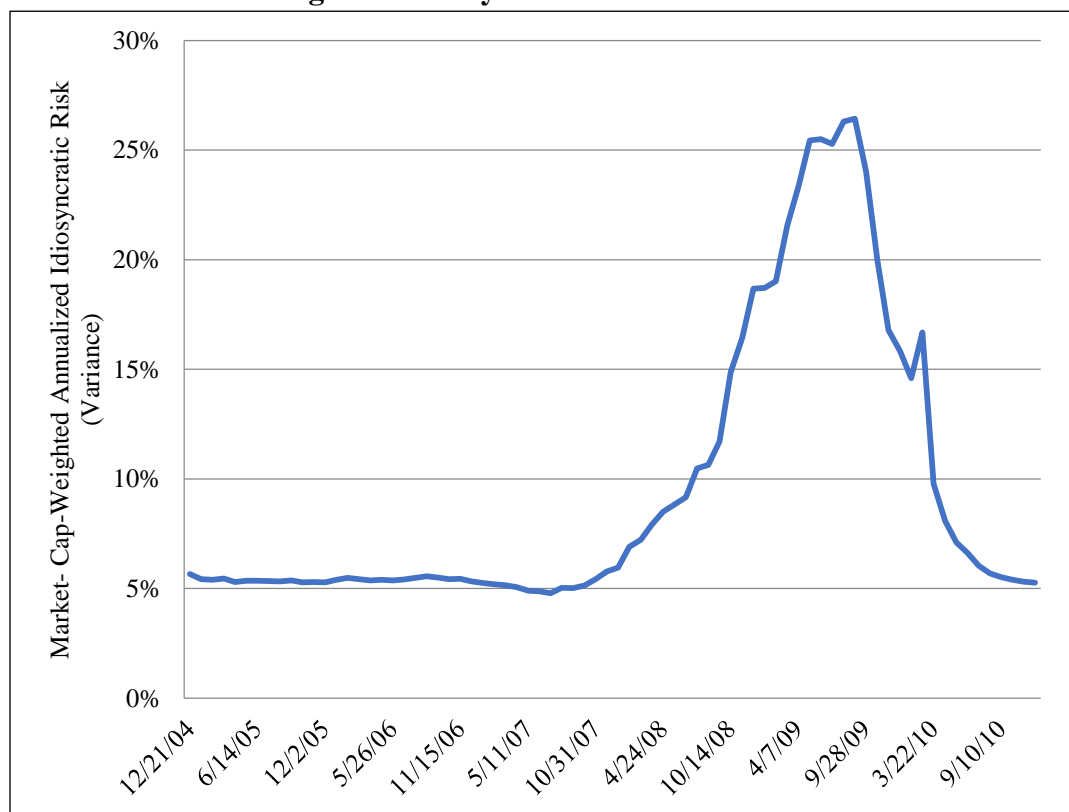
146. See Lehman Makes It Official in Overnight Chapter 11 Filing, Wall St. J.: Crisis on Wall St. Blog (Sept. 15, 2008, 7:40 AM), <http://blogs.wsj.com/wallstreetcrisis/2008/09/15/lehman-makes-it-official/> (on file with the *Columbia Law Review*) (“Lehman Brothers filed for Chapter 11 bankruptcy Monday in New York, as expected”).

147. See Edmund L. Andrews, Michael J. de la Merced & Mary Williams Walsh, Fed's \$85 Billion Loan Rescues Insurer, N.Y. Times (Sept. 17, 2008), <http://www.nytimes.com/2008/09/17/business/17insure.html?pagewanted=all> (on file with the *Columbia Law Review*) (“Fearing a financial crisis worldwide, the Federal Reserve reversed course on Tuesday and agreed to an \$85 billion bailout that would give the government control of the troubled insurance giant American International Group.”). See generally Andrew Ross Sorkin, *Too Big to Fail* (2009) (providing detailed narrative account of 2008 crisis at Lehman Brothers and AIG).

148. These were firms in the S&P 100 as of March 9, 2009. Our analysis of all firms traded on the NYSE, NASDAQ, and AMEX yielded similar results.

spike-like pattern of crisis-induced increase in idiosyncratic risk. Over a one-year period idiosyncratic risk increases sharply and then, just as sharply, returns to precrisis levels. These results are depicted graphically in Figure 5 and are reported in Table 6 below.¹⁴⁹

Figure 5: Idiosyncratic Risk 2004–2010



149. Figure 1 is calculated by first estimating the company-specific volatility for each firm that trades on the NYSE, AMEX, or NASDAQ on each day in 2004 to 2010. This volatility is found using the variance of the error term in a CAPM regression over the last year. So the volatility on January 1, 2005, represents what happened from January 1, 2004, to December 31, 2004. Likewise January 2, 2005, represents what happened from January 2, 2004, to January 1, 2005. We then average the results for each day, weighting by market capitalization of the firm. Table 6 uses the same method, but only for S&P 100 firms. In each case, this company-specific volatility is measured by using the variance of the error term in a market model regression of the firm's daily returns (percentage changes in stock price) over the last year. This regression estimates the firm's β , which captures how, on average over the year, the firm's daily stock price responds to price changes in the market as a whole. See generally Brealey, Myers & Allen, *supra* note 123, at 178–82 (“If you want to know the contribution of an individual security to the risk of a well-diversified portfolio . . . you need to measure its *market risk*, and that boils down to measuring how sensitive it is to market movements. This sensitivity is called beta (β).”). On any given day, this β can be used to estimate a predicted return for the firm based only on market performance. The error term for the day is the difference between the firm's actual return and this predicted term. The variance of this error term over all the trading days of the year is thus a measure of how much the firm's share price is being moved around by news that is independent of the news that moves the market as a whole around each day.

Table 6:
Idiosyncratic Risk of S&P 100 Firms, 2005–2009

	<i>Market-Cap-Weighted Annual Idiosyncratic Risk</i>			
<i>Period</i>	All		Financial ¹⁵⁰	Nonfinancial
July 1, 2005–June 30, 2006	3.5%		1.8%	3.8%
July 1, 2006–June 30, 2007	3.3%		1.7%	3.6%
July 1, 2007–June 30, 2008	5.7%		8.9%	5.4%
July 1, 2008–June 30, 2009	18.2%		74.0%	13.3%

B. Looking Back over Eight Decades

A relationship between downturns in GDP and idiosyncratic risk was first noted by Campbell et al. in 2001.¹⁵¹ They found that over the thirty-five-year period from 1962 to 1997, a sharp increase in idiosyncratic risk was associated with the 1970, 1974, 1980, 1982, and 1991 recessions as well as with the October 1987 market break.¹⁵² Campbell et al. did not, however, consider earlier crises, including the stock market crash of 1929 and the period of the Great Depression. Nor, of course, could they have considered the dot-com boom and bust in the early 2000s or the Great Recession of 2008 to 2009.

We have performed a study similar to Campbell et al. but extended the period covered from 1926 to the present. As depicted in Figure 6 below, the results indicate that this pattern of increased idiosyncratic risk associated with poor macroeconomic performance repeats itself throughout the much longer eighty-five-year period, with particularly high levels of idiosyncratic risk at the time of the stock market crash of 1929, the early years of the Great Depression in the early 1930s, the economy's retreat into deep recession in 1937, and the financial crisis of 2008 to

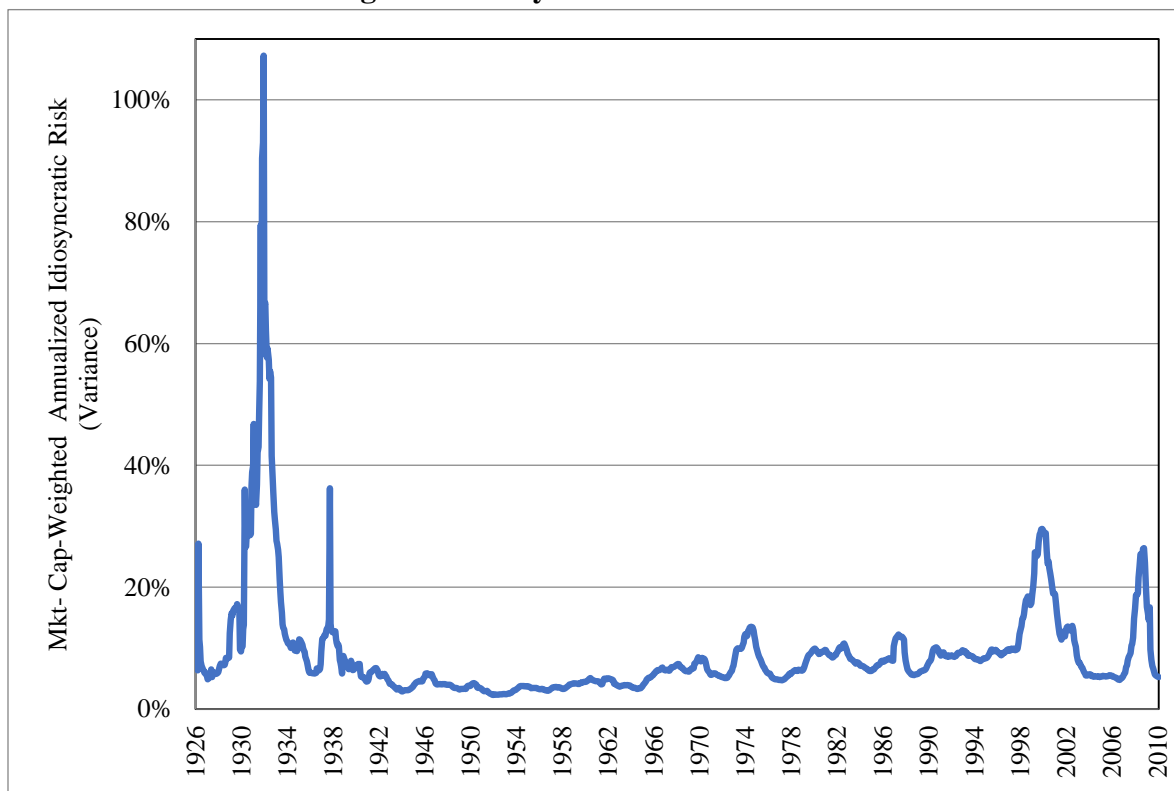
150. The companies analyzed in Table 6 were part of the S&P 100 as of March 2, 2009. By that point, the financial firms in the index were the ones who had weathered the crisis relatively well. Therefore, the exponential increase in volatility for financial firms is not attributable to companies like AIG, Bear Stearns, and Lehman.

151. Campbell et al., *Empirical Exploration*, supra note 141, at 3.

152. See id. at 13 fig.4 (showing “[a]nnualized firm-level volatility”).

2009.¹⁵³ We also find that idiosyncratic risk increases at times of market boom as well, although the relationship is weaker—a point to which Part III will return.

Figure 6: Idiosyncratic Risk 1925–2010



C. Sectoral Analysis

As shown in Table 6, the increase in financial-sector volatility during the 2008 to 2009 crisis dramatically outpaced that in the nonfinancial sector. It is therefore reasonable to ask whether the average increase among nonfinancial firms during the crisis was simply the shadow of the crisis in finance falling on a few adjacent industries such as construction. The answer is no. There was a substantial increase in average idiosyncratic volatility in each of the sixty-two-digit Standard Industrial Classification (SIC) industries surveyed.¹⁵⁴ Every industry saw its idiosyncratic

153. The results depicted in Figure 2 are obtained in the same manner as those in Figure 1. See *supra* note 149 (describing analysis). Our method is slightly different than that used by Campbell et al., but our results for the period that our study and theirs overlap are very similar. Appendix, *supra* note 144, at fig.(2)A-2.

154. Appendix, *supra* note 144, at tbl.(2)A-2.

volatility, as measured by variance, increase more than 50%, and in fifty-eight of sixty sectors it more than doubled.¹⁵⁵

Although firms in every industry experienced significant increases in idiosyncratic risk, what distinguishes those industries that experienced the greatest increase? Interestingly, the sectors whose firms would typically be seen as the riskiest prior to the crisis—those whose stock prices were most sensitive to changes in general economic conditions (i.e. those firms with the highest precrisis β (beta))¹⁵⁶—were not the sectors in which idiosyncratic risk increased the most. Rather, it was those sectors with relatively *low* precrisis β s that saw the largest increases in their idiosyncratic risk.

While a high absolute level of β prior to the crisis does not explain the impact of an economic crisis on a company's idiosyncratic risk, there is a significant relationship between crisis period *increases* in β and the industries that showed the largest increases in idiosyncratic risk.¹⁵⁷ That is, those firms whose stock prices became more sensitive during the crisis to changes in the overall economy, even though they had a low precrisis β , also tended to have the greatest increase in idiosyncratic risk.¹⁵⁸ This finding supports the first of our possible

155. Codes 60–67 are financial. The largest increases among nonfinancial firms were in hotels, amusement services, lumber, and social services. As Figure A-3 in the online appendix shows, controlling for industry-specific factors, along with those of the overall market, does not alter the results. Appendix, *supra* note 144, at fig.(2)A-3.

156. See Brealey, Myers & Allen, *supra* note 123, at 178–82 (defining “beta” as measure of “how sensitive” particular asset “is to market movements”).

157. The results noted in this paragraph are robust to the inclusion or exclusion of financial firms.

158. One question about this finding should be addressed. We measure a firm's idiosyncratic volatility by running a market model. see John Y. Campbell, Andrew W. Lo & A. Craig MacKinlay, *The Econometrics of Financial Markets* 149–80 (1997) [hereinafter Campbell, Lo & MacKinlay, *Econometrics*], to net out the impact of the broader market on the individual stock. In so doing, we estimate the firm's β , which captures how the firm's stock responds to changes in the market. *Supra* note 149. Using this β allows one to estimate a predicted return for the firm based only on market performance. The idiosyncratic risk is derived from how much the actual return varies from the predicted return. If the estimated β departs from the “true” β , the predicted return will be less accurate. This will increase measured idiosyncratic risk regardless of whether the estimated β is too large (in which case when the market goes up, the predicted return will be too positive and if the market goes down, too negative) or whether β is too small (in which case when the market goes up, the predicted return will not be positive enough and when the market goes down, it will not be negative enough).

explanations for the reported crisis-related increase in unsystematic risk: When a firm is subject to a structural change in the economy, as reflected by an increase in the firm's β , each new bit of firm-specific information will take on greater significance and hence move price more.

D. The Effects of Leverage

An alternative explanation for all or part of our reported increase in firm-specific volatility during economic crises is that the increase is caused by an increase in firms' leverage. Because the value of most firms' equity decreases during crises, their debt–equity ratio increases correspondingly. The 2008 to 2009 financial crisis illustrates the phenomenon: From peak to trough, the market capitalization of all firms in the Center for Research in Security Prices (CRSP) database declined 55%. It is well understood that an increase in a firm's debt–equity ratio increases the variability of both the systematic and, of particular interest to us, the idiosyncratic portions of a firm's returns.¹⁵⁹ This is because debt is paid first from the firm's underlying cash

As detailed in section I.D and in the online appendix, for the typical firm this estimate of β , though unbiased, is less accurate during crisis times than in normal times because the actual value of β over the period of the year is likely to change more in crisis times due to changes in the firm's leverage triggered by sharp changes in the value of its equity. Appendix, *supra* note 144, at 8–11. Thus, one might speculate that the reported result—that firms in sectors whose β s, as we estimated them, increased the most also had the largest increases in idiosyncratic risk—might be an artifact of this less accurate assessment of their β s rather than evidence of a positive relationship between an increase in a firm's true β during crisis and firm-specific risk. The underlying intuition would be that most of the firms whose β s appeared to have experienced the greatest increases might in fact be those whose estimated β s, by chance, deviated the most on the upside from their actual β s. If this is correct, then our assessments of their idiosyncratic risk would also be more inflated than firms where the estimated β was closer to the true value.

This concern, however, appears to be unwarranted. If it were correct, we should, for the same reasons of chance, also see an increase in firm-specific risk in sectors whose estimated β s dropped dramatically during the crisis because they would have a concentration of firms whose estimated β s deviated the most on the downside from their true values. In fact, sectors whose estimated β s dropped during the crisis were no more likely to see a relatively large increase in firm-specific risk than those whose estimated β s were unchanged.

As a further sensitivity check, we also measured firm-specific risk using very short periods (twenty trading days) to minimize the possible impact of the decline in the accuracy of our β estimates compared to those based on a full year, during which the value of the true β may have shifted much more substantially than in a month. The results, reported in the appendix, are very similar to those presented in the text, again suggesting that the decline in the accuracy of our β estimates does not explain the results presented in the text. *Id.* at 12.

159. See Brealey, Myers & Allen, *supra* note 123, at 434–35 (“The expected rate of return on the common stock of a levered firm increases in proportion to the debt-equity ratio”); Franco Modigliani & Merton H. Miller, *The Cost of Capital, Corporation Finance and the Theory of Investment*, 48 *Am. Econ. Rev.* 261, 267–72 (1958) (“[T]he expected yield of a share of stock is equal to the appropriate capitalization rate p_k for a pure equity stream in the class, plus a premium related to financial risk equal to the debt-to-equity ratio times the spread between p_k and r .” (emphasis omitted)).

flows, with equity receiving whatever residuals there are. Absent insolvency, the residual going to equity thus bears all the variability in the firm's cash flow. When the total value of the equity in a firm is reduced, this variability will be concentrated on fewer dollars of equity and so the return on each such dollar varies more.

Perhaps, then, the crisis-related spikes in idiosyncratic risk simply reflect the effect of the crisis-related decreases in equity value.¹⁶⁰ Then there is no puzzle to explain, and the focus can turn directly to the legal implications of these findings. To address this possibility, we added explicit controls for the effect on idiosyncratic risk of the increase in leverage.¹⁶¹ As depicted in Figure 7, however, no more than one-quarter of the observed increase in idiosyncratic volatility in the most recent crisis was due to the leverage effect that arises from depressed share prices.¹⁶²

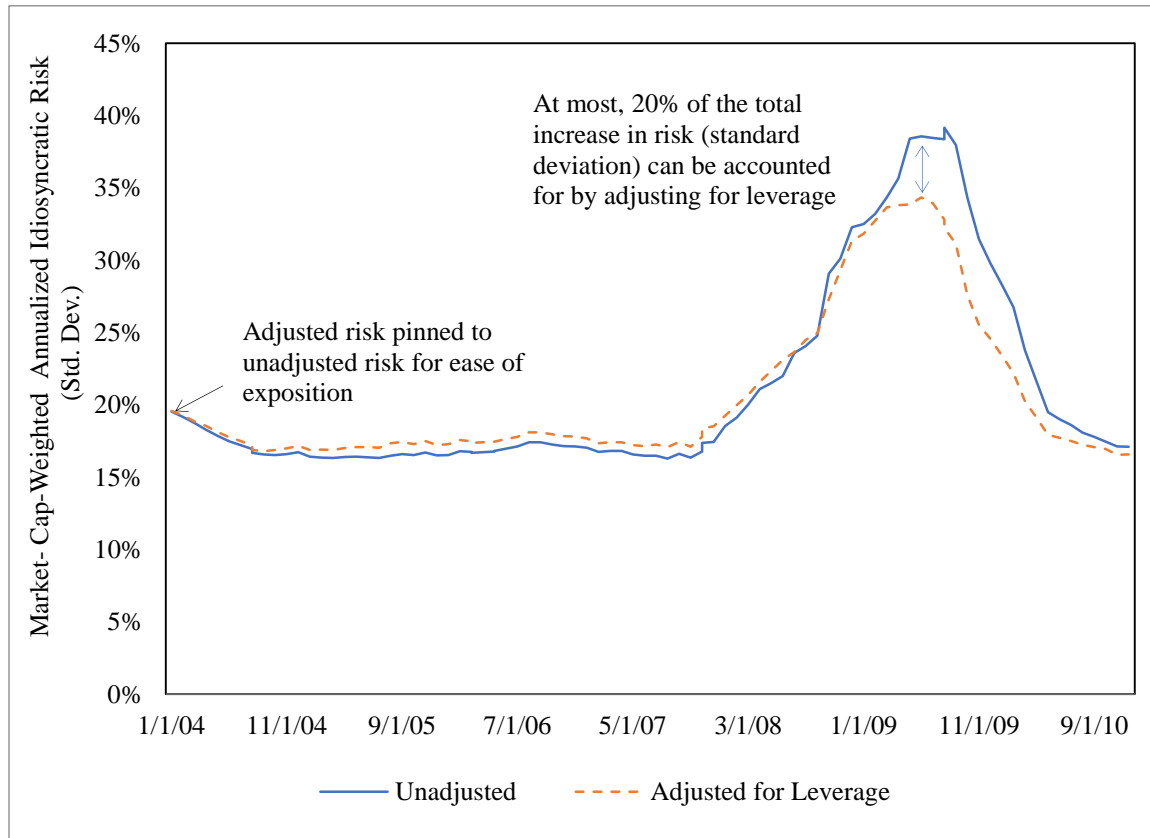
Thus, the puzzle with which this Part began still remains: Why, during periods of economic crisis, are large increases in idiosyncratic risk consistently associated with a company's stock? We address this puzzle in the next Part.

160. We are grateful to Mark Weinstein for raising this point.

161. For each firm, we adjust for the effect of leverage by deflating the unadjusted daily volatility by multiplying it by the firm's Equity/(Debt+Equity) on that day. For more details, see Appendix, *supra* note 144, at Fig. (2)A-6 (explaining mathematical computation of unlevered β based on firm's ratio of debt to equity).

162. The online appendix also discusses and seeks to control for a second leverage-related impact on our measure of idiosyncratic risk, a changing ("unstable") β during the observation period. *Id.* at 11. This too will increase the level of idiosyncratic risk measured for the typical firm. We find, however, that the second effect is in fact even less important than the first and so explains relatively little of the crisis-induced increase in idiosyncratic risk that we have observed.

FIGURE 7: IDIOSYNCRATIC RISK OF S&P 100 FIRMS ADJUSTED FOR LEVERAGE



II. POSSIBLE EXPLANATIONS

As noted above, Campbell et al. first reported that idiosyncratic risk increases during crisis periods.¹⁶³ Their focus, however, was not on explaining this link, but on explaining what they saw as a secular increase in idiosyncratic volatility from 1962 through 1997. We focus here on the puzzle presented by extending their observation period to the eighty-five years from 1926 through 2010. From the lead up to the Great Depression through what has come to be called the Great Recession, economic downturns repeatedly coincide with spikes in idiosyncratic volatility. Why should a crisis be associated with an increase in risk unrelated to the general economic disorder? We consider three possible explanations. One, explored in section II.A is that, compared to ordinary times, information about a firm contained in current news may become

163. Campbell et al., *Empirical Exploration*, supra note 141, at 3.

more important in predicting its future cash flows relative to the role of the already existing stock of knowledge in making such predictions. A second explanation, explored in section II.B, is that the quality of management becomes more important in crisis times. Consequently, the ordinary flow of new information about this subject can cause bigger movements in price because each bit tells the market about something—the quality of management—that the market now regards as more important than it did in ordinary times. A third explanation, explored in section II.C, is that crisis creates uncertainty as to what factors, and hence what information, are important to valuation. Because of this uncertainty, a broader range of information has valuation implications and therefore stock prices move more frequently. These three explanations share a common theme: Economic crises increase both the frequency and the impact of news with valuation consequences for an individual company's stock. New information becomes more valuable and more information is likely to be new, with the consequence of increased idiosyncratic volatility.

A. *Current News Becomes Relatively More Important*

In a rational market, share price at any given moment reflects the aggregate of investors' predictions of an issuer's future net cash flows.¹⁶⁴ These predictions are based on a large collection of bits of information, much of it accumulated over a period of years, but some of which at any time is new. In an informationally efficient market, the new information in each bit is by definition unpredictable before it is received. This new information is what causes the issuer's share price to fluctuate in a random walk: If new information cannot be predicted, then neither can the direction of its impact on price. How much a stock's price moves one way or the

164. See Brealey, Myers & Allen, *supra* note 123, at 80 (noting present value of share of stock equals expected future dividends per share, discounted to present value).

other depends on the predictive importance of the new bit of information relative to the predictive value of the previously accumulated information concerning that stock.

A first explanation for the association between economic crises and spikes in idiosyncratic volatility is that during an economic crisis, each bit of new information is likely to be more important relative to the existing stock of information for predicting the future. This should be true not only for information relevant to predicting the economy's overall performance but also for information that is relevant to predicting the future cash flows of only a single firm.

One reason that the predictive importance of firm-specific new information increases in an economic downturn is that investors expect structural changes in the economy to accompany a downturn. This is at least in part true because major downturns are usually the result of imbalances in the economy: in the case of the recent financial crisis, an unsustainable level of resources going into construction of residential and commercial buildings and into the financial industry. But the exact nature of this structural change and its implications, good or bad, for any particular firm might not yet be fully understood, as the deluge of books seeking to explain the Great Recession strikingly demonstrates.¹⁶⁵ In essence, an economic crisis “shakes the box.” Everyone knows that when things settle down, the relations among firms, and those between each firm and its suppliers and customers, are likely to be different than before the crisis. Relative to each other, some firms will gain and others will lose from these changed relations. But no one is yet sure exactly what pattern these new relations will take and hence what companies will be the winners or losers. A new bit of firm-specific information can have a bigger

165. See Gary Gorton & Andrew Metrick, *Getting up to Speed on the Financial Crisis: A One-Weekend-Reader's Guide*, 50 *J. Econ. Literature* 128, 128–29 (2012) (proposing “reading list” of sixteen key documents for understanding crisis); Lo, *supra* note 127, at 154 (reviewing twenty-one books about crisis).

price impact than in normal times because the bit may disclose something about this new pattern of relations.

Consider, for example, a not-fully-anticipated quarterly earnings announcement that is 10% lower than would have been predicted by the earnings path over the previous few years. This shortfall could be due to random firm-specific factors not likely to repeat themselves, or it could be due to a more enduring change in the fundamental factors shaping the environment in which the firm operates—the costs of its inputs, the market for its outputs, or the technology that the firm uses to transform one into the other—that will continue to influence the firm's cash flows for many periods to come. A not-fully-anticipated quarterly earnings shortfall typically would have some impact on investor predictions about a firm's future cash flows and hence on its share price. In making these predictions in normal times, however, investors might assign much more weight to the firm's longer-term history of earnings because they are indicative of the environment within which the firm had been operating and there would be little reason to believe that the environment had changed radically during this most recent quarter.

During an economic crisis, in contrast, changes in the structure of the economy that significantly impact the firm are much more likely. However, there will at first be no clear understanding of the nature of these changes. The new piece of information concerning the quarterly earnings shortfall thus takes on more importance because there is a greater likelihood that the change in earnings is due to changes in more enduring factors affecting the long-run success of the firm, rather than to fluke factors unlikely to repeat themselves. Put differently, in

such times, new information takes on greater importance because the economic crisis degrades the value of the old stock of information in helping to predict the future.¹⁶⁶

This explanation is consistent with a second feature of our results. As shown in Figure 5, the crisis-induced increase in idiosyncratic risk is truly spike-like: The increased volatility disappears as quickly as it appears. This suggests that once the market figures out the new shape of the postcrisis structural relationships, idiosyncratic risk returns to historical levels. In this explanation, new information is important to understanding the crisis-induced change in the structure of relationships among firms, and the volatility spike is the artifact of that importance. Once it is understood, the cause of increased idiosyncratic volatility disappears because new information no longer carries the extra, crisis-related capacity to cause large changes in predictions.

The structural change explanation for the increase in idiosyncratic volatility during crisis periods is also supported by the finding that firms in sectors that experienced the largest increase in their β s during the recent financial crisis—that is, ones that relative to other firms became more sensitive to factors that affect the market as a whole—also displayed the greatest increases in idiosyncratic risk.¹⁶⁷ It is reasonable to think that firms that were most affected by crisis-induced changes in the structure of the economy would tend to experience the biggest changes in their β s. These are the firms for whom, during a crisis, new bits of information would have the most predictive power relative to the prior stock of knowledge. In accordance with our first

166. It should be noted that, even in normal times, a relatively small earnings shortfall from what was expected can have a significant effect on price. The earnings miss may indicate that despite all efforts to massage the numbers, the company still could not make the estimate, which suggests that something more important than a small miss has occurred. In crisis times, though, the shortfall would still, for the reasons discussed in the text, be expected to lead to an even bigger price decline.

167. See *supra* section I.C (presenting sectoral-analysis findings).

explanation, these are the firms that would display the greatest increases in idiosyncratic volatility.

A period where the market is rising rapidly, for example during the Internet boom, also may signal disruption and structural change, where again new information takes on more importance. Our findings are consistent with this circumstance as well—idiosyncratic risk goes up in sharp market upturns.¹⁶⁸ The increase is not as great as in market downturns, but the smaller impact is likely at least in part due to the deleveraging effect of rising stock prices, the opposite of the leverage effect we analyzed in section I.D.

The following hypothetical demonstrates the first explanation. Imagine that you have a barrel with 200 colored balls in it, divided in a ratio between red and green that you are trying to estimate. Each period, you randomly draw one ball, note its color, and put the ball back in the barrel. After, say, twenty periods, you will have a pretty good sense of the ratio in the barrel, and you will not change your estimate very much based on whether the twenty-first ball drawn is red or green. Now suppose that of the 200, 100 randomly selected balls are taken out of the barrel and 100 new balls are substituted for them. The newly substituted balls have an unknown ratio of red to green that might be quite different than the ratio of the 200 original balls. When you take out the twenty-second ball, the outcome (that is, the new bit of information)—whether it is red or green—will change your estimate of the ratio in the barrel much more than it would have absent the substitution.

B. Information Concerning Quality of Management Becomes More Important

The phenomenon of crisis-enhanced new information has a special role when the information concerns the quality of management. When troubled times sharply change the

168. Supra Figure 2.

overall business environment, the importance of an established firm's management becomes more like the importance of a startup's management in normal times. Management confronts more, and more important, decisions than in normal times.

It is well recognized that startup management has special importance because the company's value is primarily comprised of future growth options, as opposed to being based on the cash flows from an existing business.¹⁶⁹ At this early stage, management needs to make more choices of the kind that shape the fundamental direction of the firm. Moreover, the consequences of these choices are less obvious than they are with an established firm in normal times because there is no history of the consequences of making similar decisions in the past. Having high-quality managers—persons who are better at predicting the consequences of their more difficult choices despite the lack of such history—will thus have a larger impact on expected performance. Crisis thus can put the managers of even mature firms in a situation closer to that of a startup. The sharp change in the mature firm's business environment may force managers to make more fundamental choices, and history will be a less reliable guide as to their consequences. As a result, the quality of management has a larger impact on the firm's future cash flows.

In addition to the extent to which each bit of crisis-related information newly reveals something about the quality of a mature firm's management, the subject that bit of new information concerns—management quality—also has greater implications for the firm's future cash flows. Therefore, the significance of the bit in predicting these future cash flows becomes more important. As a consequence, the revelation of new information concerning management

169. See, e.g., Ronald J. Gilson, *Engineering a Venture Capital Market: Lessons from the American Experience*, 55 *Stan. L. Rev.* 1067, 1076–77 (2003) (noting at start-up stage, “quality of the company's management . . . takes on heightened importance because so large a portion of the portfolio company's value depends on management's future decisions”).

quality causes a bigger share price change and therefore, greater volatility than in noncrisis times.

C. Model Failure: Ignorance Concerning What Facts Are Initially Relevant

Investors' expectations about a company's future cash flows are typically based on an implicit model that tells them something about the meaning and importance of new information about the company. In essence, investors have a model that tells them what information is important for predicting future cash flows. In an economic crisis, systemic changes may make investors less confident that they still know which facts have valuation implications. Because of that uncertainty, a much wider range of facts potentially matter and so the appearance of more kinds of information will move prices. Moreover, any particular new bit of information can have greater importance. Not only does it directly tell investors something more about the firm's future cash flow, it may also provide information about the shape of a new valuation model that suggests a different range of information that will be relevant to predicting the firm's cash flows in the future. This may put an already existing stock of knowledge in new perspective, in effect, turning old information into new. Thus, new information and more kinds of information can potentially move price a great deal.

This third explanation differs from the first explanation—that current news becomes more relevant—in that the situation in the first explanation could be fully described in terms of risk, while the third explanation resembles more Knightian uncertainty.¹⁷⁰ Continuing with the analogy of sampling from a barrel of balls, you still need to estimate the ratio of the colors of the balls in the barrel, but you can no longer even be sure that there are only red balls and green balls

170. See Frank H. Knight, *Risk, Uncertainty and Profit* 233–34 (1921) (distinguishing between risk, which can be predicted probabilistically, and uncertainty, which cannot).

after the substitution of the 100 new balls. Suppose, after this substitution, the twenty-second ball drawn is yellow, not red or green. The draw tells you not only something about the ratio of balls, it tells you that your old model—a two-color distribution—is no longer accurate. Thus, you get a whole new kind of information relevant to predicting the contents of the barrel.¹⁷¹

III. THE EFFECTS OF CRISIS-INDUCED SPIKES IN IDIOSYNCRATIC VOLATILITY ON EVENT STUDIES

Over the last few decades, financial economics and related econometric tools have become increasingly integrated into corporate and securities law analysis and practice. The synergy is straightforward: Financial economics speaks to how assets are valued and corporate and securities laws provide a structure in which value can be created. This sensible trend has proceeded, however, without an appreciation of the dramatic several-fold increase in idiosyncratic volatility accompanying economic bad times that we have documented here and sought to explain. Yet the sharp price drops that accompany bad times make these times precisely when there are more plaintiffs' suits, more governmental enforcement actions, and more calls for regulatory change. This Part and Part IV take a fresh look at the link between financial economics and corporate and securities law in light of what we have shown is the predictable

171. Warren Buffett has famously made a related point. In his terms, “[Y]ou only find out who is swimming naked when the tide goes out.” Letter from Warren E. Buffett, Chairman, Berkshire Hathaway Inc., to Shareholders of Berkshire Hathaway Inc. 10 (Feb. 28, 2002), <http://www.berkshirehathaway.com/letters/2001pdf.pdf> [<http://perma.cc/QUJ7-X8PG>].

Put more precisely, but less amusingly, good times allow a pooling equilibrium concerning firm quality, and the economic crisis results in separation. While Buffet’s point covers part of our analysis, it does not address an important aspect of our data. Our measure documenting the idiosyncratic volatility increase involves the variance of daily market-adjusted price changes over a twelve-month period. For most firms in the sample, much of the increased variance in the idiosyncratic portion of the firm’s returns comes from an increase in both the upside and the downside. In contrast, Buffet’s separating effect of bad economic times—revealing which are the bad firms and which are the good ones—would, for any given firm, work in just one direction. Moreover, if the large declines in share price associated with the revelation of dishonest or incompetent management drove a significant portion of the crisis-times increase in idiosyncratic risk, then, compared to normal times, the largest drops should explain a substantially higher portion (relative to what we have observed) of the total amount of such risk during the twelve-month period. As discussed in the online appendix, in crisis and in precrisis times alike, about the same portion of total idiosyncratic risk is explained by the biggest drops, a result inconsistent with the “tide goes out” explanation. See Appendix, *supra* note 144, at 8 fig.(2)A-5 (comparing unadjusted market-cap-weighted volatility with estimates after truncating tails).

spike in idiosyncratic volatility that accompanies each major economic downturn. Sections IV.A and IV.B consider the impact of crisis-related spikes in idiosyncratic volatility on the use of event studies in connection with a number of issues that arise in securities litigation. Section IV.C then turns to the impact of these spikes on the extent to which incumbent management's valuation of a target company should receive deference in the context of a contest for corporate control.

Financial economics focuses centrally on what factors influence the price of a security. The inquiry can be framed by identifying an event that may influence a security's price and then seeking to measure the impact of that event on the price of the security in question. Measuring the price impact of a particular event, however, requires isolating the effect of a single item of information—the occurrence of the event under study—from the cacophony of information constantly reaching the capital markets. This is the province of an event study.¹⁷² Over the last few decades courts have come, as a practical (and probably as a legal) matter, to require plaintiffs to conduct an event study for many securities fraud class actions to proceed.¹⁷³ As we will see, the level of idiosyncratic volatility associated with a company's stock is central to the event study methodology and the sharp spikes in such volatility that accompany economic crises cause problems. To see why, we need to start with a brief account of how an event study is conducted.

172. See, e.g., Sanjai Bhagat & Roberta Romano, Event Studies and the Law Part I: Technique and Corporate Litigation, 4 Am. L. & Econ. Rev. 141, 142 (2002) (explaining methodology of event studies focuses on movement of stock prices due to specific events); Sanjai Bhagat & Roberta Romano, Event Studies and the Law Part II: Empirical Studies of Corporate Law, 4 Am. L. & Econ. Rev. 380, 380 (2002) (“[E]vent studies provide a metric for measurement of the impact upon stock prices of policy decisions.”).

173. See, e.g., *Bricklayers & Trowel Trades Int’l Pension Fund v. Credit Suisse First Bos.*, 853 F. Supp. 2d 181, 186 (D. Mass. 2012) (“An event study . . . often plays a ‘pivotal’ role in proving loss causation and damages in a securities fraud case. Given the difficulty inherent in proving the effect, if any, of a single news item on the price of a stock, many courts require them in such cases.” (citation omitted)); *In re Williams Sec. Litig.*, 496 F. Supp. 2d 1195, 1272–73 (N.D. Okla. 2007) (“[A] number of courts have rejected or refused to admit . . . damages reports or testimony by damages experts in securities cases which fail to include event studies or something similar.” (quoting *In re Imperial Credit Indus., Inc. Sec. Litig.*, 252 F. Supp. 2d 1005, 1015 (C.D. Cal. 2003))).

A. *Event Study Basics*

It would be very useful to be able to identify the impact of a particular event on a stock's price. For example, on the day that any particular item of news relating to a specific issuer becomes public, say the announcement of an acquisition, other bits of news concerning the company also may affect its share price. So the mere fact that the share price moved up or down that day does not show that the price movement was caused by the acquisition. In assessing whether the acquisition will increase a firm's value, it would be helpful to observe just the effect of the acquisition announcement—to separate the effect of the “event” from that of any other bits of news relevant to the company's future cash flows and of random noise. An event study helps distinguish between the different possible influences on stock price in order to assess the extent to which a particular item of information affected price.¹⁷⁴

1. Determining the Market-Adjusted Price Change. — Conducting an event study begins with measuring the market-adjusted change in the issuer's share price when the item of interest becomes public. The market-adjusted change is the difference between the observed price change and what the change in overall stock market prices predicts would have been the issuer's price change. This prediction is based on the historical relationship (usually over a one-year observation period ending shortly before disclosure of the item of interest) between price changes in the overall market and price changes of the issuer's stock.

174. Those familiar with event studies will recognize a fundamental difference between event studies as originally developed in financial economics and event studies as they are used in securities litigation. In financial economics, the key to an event study is that the task is to measure the effect of an event in general—does it affect stock price over a sample of companies that experienced the event?—not the effect of an event on a particular company's stock price. The financial economics approach uses a sample of all companies that experienced the event in question to regress out the effect of all the other events affecting a company's stock price. If the number of companies in the sample is large enough, the other events affecting particular companies cancel each other out. See Campbell, Lo & MacKinlay, *Econometrics*, supra note 158, at 149–80 (explaining event study methodology).

A single-company event study, as used in the securities litigation context, must address the problem of other events occurring close in time to the event under study in other ways. The text is concerned with single-firm event studies.

Making this market adjustment removes the influence on the stock's price of news that affects the price of all firms' stock, i.e., bits of systematic news. What is left—the market-adjusted price change—is the portion of the observed change in price that is due to firm-specific news or background noise. Because in securities litigation, the news item of interest relates specifically to the issuer, eliminating the impact of systematic news is critical.¹⁷⁵

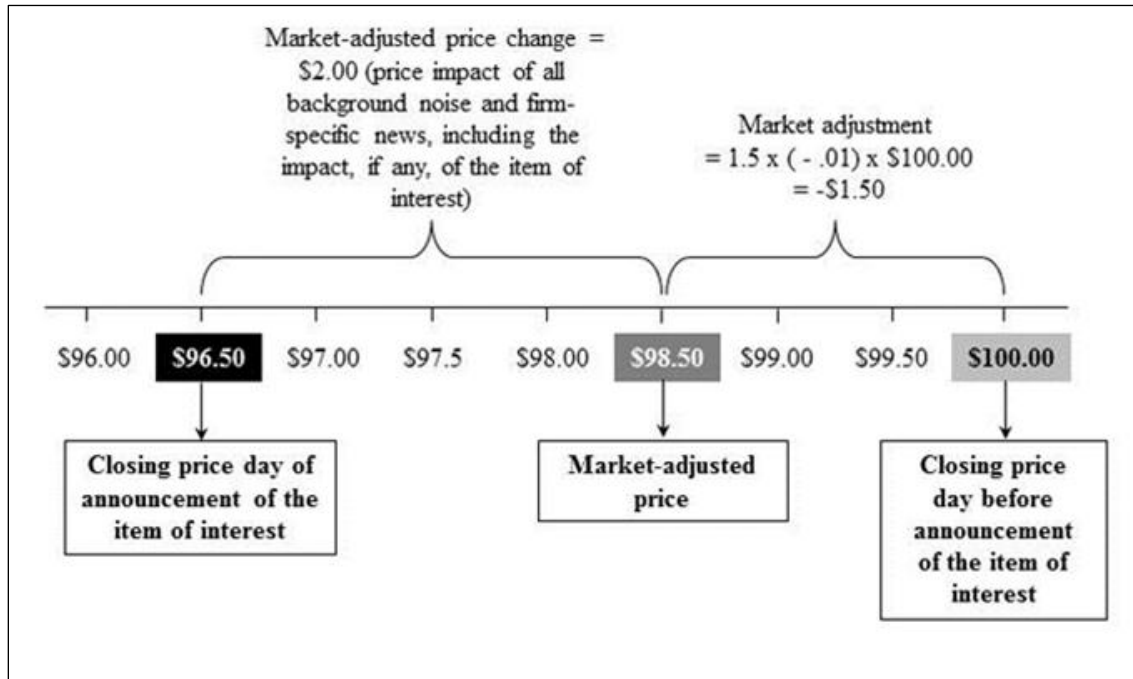
To give an example, suppose that the share price of the issuer under study was \$100.00 at the end of the trading day immediately preceding the item of interest's disclosure and is \$96.50 by the end of the day on which the item of interest is disclosed. Suppose as well that the market as a whole went down 1% on the day of the item of interest and that the issuer's β is 1.5.¹⁷⁶ Based on this historical relationship between day-to-day changes in the issuer's share price and the corresponding market-wide price changes, we would predict that if firm-specific news, including the item of interest, had on a net basis no effect on the issuer's share price, the issuer's price would have dropped to \$98.50. But in fact it dropped to \$96.50. So the remainder of the observed price change—referred to as the market-adjusted price change—would be $-\$2.00$, or -2.00% . As depicted in Figure 8, this is the portion of the total observed price change that can be attributed to firm-specific news.¹⁷⁷

175. Courts recognize the need to separate systematic from unsystematic movements in stock price based on the utility of single-firm event studies in accomplishing this. For examples of courts requiring the use of event studies to strip away any movement in price caused by market-wide trends for plaintiffs to establish loss causation, see *Imperial Credit Indus., Inc.*, 252 F. Supp. 2d at 1015–16 (rejecting testimony of plaintiff's expert for failure to include "event study or similar analysis . . . [to] eliminate that portion of the price decline . . . which is unrelated to the alleged wrong," but attributable to "market events for which Defendants cannot be held responsible" (citations omitted)); *In re N. Telecom Ltd. Sec. Litig.*, 116 F. Supp. 2d 446, 460 (S.D.N.Y. 2000) ("[Plaintiff's expert's] testimony is fatally deficient in that he did not perform an event study or similar analysis to remove the effects on stock price of market and industry information . . .").

176. See *supra* note 149 (explaining concept of β).

177. See Campbell, Lo & MacKinlay, *Econometrics*, *supra* note 158, at 149–80 (discussing methodology of event studies).

Figure 8: Isolation of Firm-Specific News Through Market-Adjustment



2. *Judging the Market-Adjusted Price Change Against Its Historical Volatility.* — The next step in an event study determines the likelihood that at least part of the observed market-adjusted price change results from the item of interest. In essence we are asking the following question: How likely is it that we would observe a market-adjusted price change of the magnitude that was observed on the day the item of interest was announced if in fact the change results solely from the day's other bits of firm-specific news and background noise? This can be answered by comparing the issuer's market-adjusted price change on the relevant date with the historical record of the daily, market-adjusted ups and downs in the issuer's share price, typically over the approximately 250 trading days in a one-year observation period ending on a day shortly before the item of interest's disclosure, i.e., by comparing the magnitude of the market-adjusted price change on the day of the announcement with the issuer's historical idiosyncratic volatility.

As a general matter, market-adjusted price changes, up and down, are distributed in a pattern closely resembling what would be produced by a normal (bell-shaped curve) probability

distribution with a zero mean.¹⁷⁸ The conventional event study assumes that the same probability distribution generates the market-adjusted price change on each of the approximately 250 trading days during the observation period *and* the net price impact of all ordinary, day-to-day firm-specific news and background noise other than any impact from the item of interest on the day that it is announced. Because of the size of the sample, the standard deviation of the observation period's approximately 250 market-adjusted price changes is a fairly precise estimate of the standard deviation of this probability distribution.

Because this probability function is a normal distribution, the net price impact of all the other firm-specific news of the day and background noise will be within plus or minus 1.96 standard deviations of the mean, 95% of the time. From this, we can see that if the item of interest in fact had no impact on price, there is less than a 5% chance that we would observe a market-adjusted price change on the day it was announced that is plus or minus, more than 1.96 standard deviations from the mean. Therefore, if we observe a market-adjusted price change this large, we can reject with at least 95% confidence the null hypothesis that the item of interest had no impact on price. Accordingly, observed market-adjusted price changes that are large enough to pass this test are often referred to as being statistically significant at the 95% level.¹⁷⁹

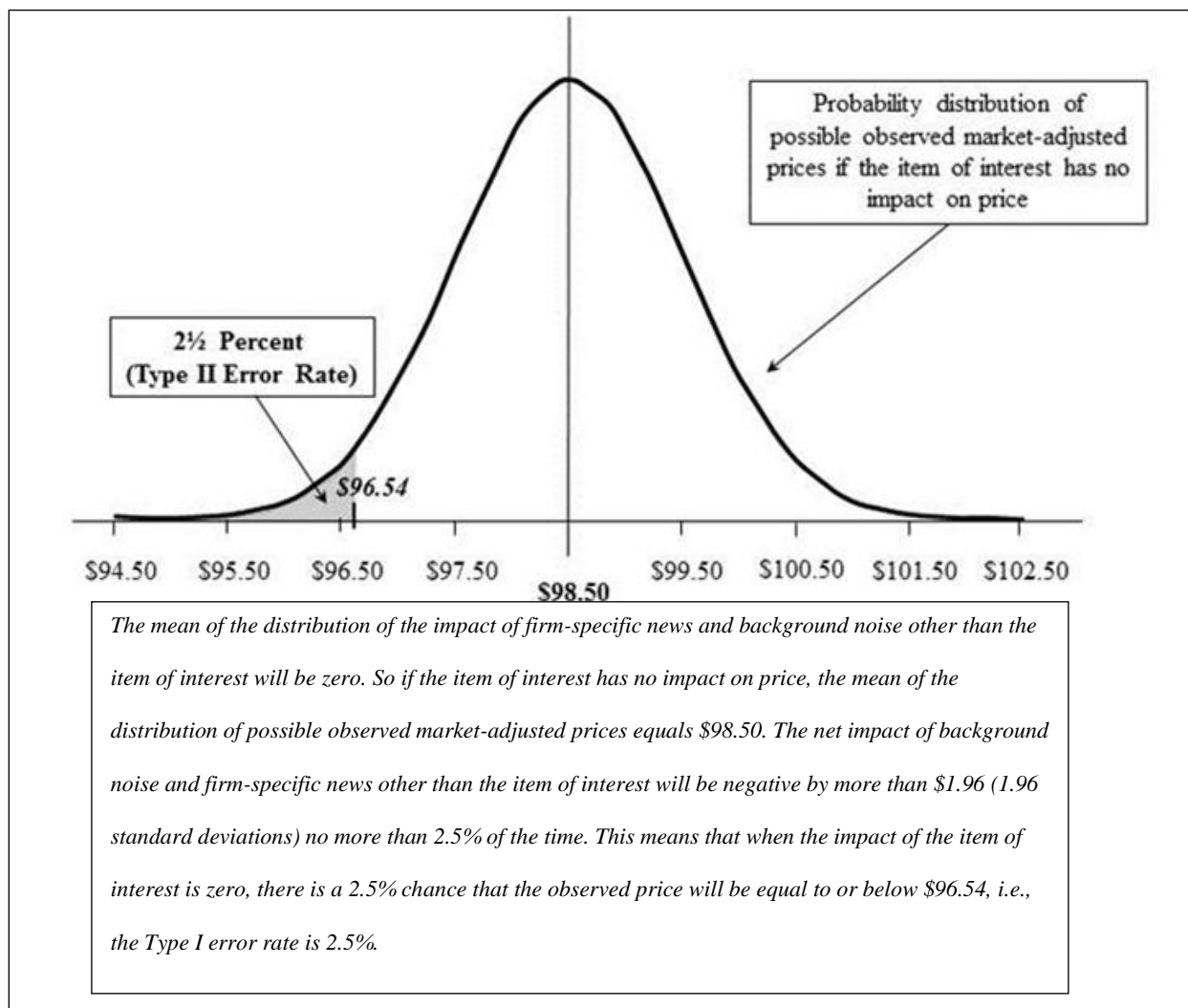
178. While the event studies used by experts in securities litigation cases almost universally assume that these price changes are normally distributed, it has been recognized for some time that the actual distribution of these changes is not perfectly normal. See, e.g., Stephen J. Brown & Jerold B. Warner, Using Daily Stock Returns: The Case of Event Studies, 14 J. Fin. Econ. 3, 4–5 (1985) (discussing non-normal distribution in daily stock returns). This has led some commentators to call for using another technique for conducting event studies in securities litigation situations. See, e.g., Jonah B. Gelbach, Eric Helland & Jonathan Klick, Valid Inference in Single-Firm, Single-Event Studies, 15 Am. L. & Econ. Rev. 495, 496–99 (2013) (explaining flaws of standard approach and proposing alternative test). The overall analysis in this Article would apply equally to these other techniques.

179. See generally Campbell, Lo & MacKinlay, *Econometrics*, supra note 158, at 149–80 (explaining history and methodology of event study analysis). As a technical matter, saying that we can reject the null hypothesis with “95% confidence” is the province of Bayesian statistics. Our terminology, however, is commonly used with event studies and shall be thought of as shorthand for saying there is a less than 5% chance of observing a result this extreme if the null hypothesis is true.

These points may be most easily understood by extending our example above. Suppose, for purpose of illustration, that the standard deviation of the market-adjusted price changes during the observation period was 1%. This would mean that on the day the item of interest is announced, the net price impact of the other firm-specific news and background noise would 95% of the time be somewhere between +1.96% and -1.96%, as illustrated in Figure 9. It would be outside this range on the positive side 2.5% of the time and on the negative side 2.5% of the time. In our example, the observed market-adjusted price change on the day of the announcement of the item of interest was -\$2.00 or -2.00%. This is more than 1.96 times the standard deviation of the day-to-day ups and downs in the market-adjusted price during the observation period, and so the price change was statistically significant at the 95% level.

FIGURE 9: DEMONSTRATION OF TYPE I ERROR

AN ANNOUNCEMENT OF AN ITEM OF INTEREST WITH NO NEGATIVE IMPACT ON PRICE WOULD BE ACCOMPANIED BY A MARKET-ADJUSTED PRICE CHANGE AS NEGATIVE AS $-\$2.00$ LESS THAN 2.5% OF THE TIME



If the difference between the issuer's market-adjusted price change when the event of interest is disclosed and its historic market-adjusted price change over the observation period is large enough to pass this test of statistical significance and no other important bits of firm-specific information become public close in time to the event of interest, we can, with at least 95% confidence, reject the null hypothesis that the observed market-adjusted price change was

due entirely to factors other than disclosure of the item of interest. In other words, the observer can reject with this level of confidence the proposition that the item of interest had no effect on price.¹⁸⁰

B. Type I and Type II Errors in the Use of Event Studies for Securities Litigation

A securities fraud class action typically is based on the theory that a corporate disclosure caused the issuer's stock price to move in a particular direction. An event study is used to test that theory. For example, if the purchaser of a security claimed that a corporate misstatement inflated the price she paid and that she suffered a loss when the stock price dropped in response to the truth coming out, she would seek to prove that the corrective disclosure negatively affected the price.¹⁸¹

As just discussed, because there are potentially many other bits of firm-specific news and general background noise affecting an issuer's share price on the same day that the item of interest is announced, we cannot determine with certainty whether the item of interest had any negative impact on price. We instead use an event study to make a probabilistic assessment of whether the item in fact had an effect on the company's share price. An event study addresses this question by providing a probabilistic assessment of whether the corrective disclosure had a

180. The foregoing discussion presents a somewhat simplified version of the event studies that are typically conducted by experts giving testimony in securities actions. For example, there is usually a control for industry-specific pieces of information as well as for systematic information, which will entail some complications concerning the proper definition of the industry. Also, sometimes the event window in which the market-adjusted price change is measured is longer than one day. In addition, often the baseline one-year observation period used to determine the standard deviation has removed days on which there are identifiable, obviously important bits of firm-specific news and in parallel, there is an attempt to remove the impact of any other identifiable, obviously important bits of firm-specific news on the day that the item of interest is disclosed. For a discussion of these elaborations on the simple model presented here, see, e.g., Allen Ferrell & Atanu Saha, *The Loss Causation Requirement for Rule 10b-5 Causes of Action: The Implications of Dura Pharmaceuticals, Inc. v. Broudo*, 63 Bus. Law. 163, 166–67 (2007). None of these elaborations of event-study methodology affect the points that we seek to make.

181. For purposes of illustration, we will assume throughout this Article that the legally relevant question is whether an item of news had a negative effect on price, but a symmetrical version of the discussion would be equally valid where the legally relevant question is whether the item had a positive effect on price—for example, in a case claiming that a proposed acquisition was favorable to shareholders.

negative effect on the company's share price. As the question is usually put, did the issuer's stock price decline by a statistically significant amount? For our purposes here, the critical fact is that this test will generate a certain rate of false positives (Type I errors) and of false negatives (Type II errors). Type I errors occur when the item of interest did *not* in fact have an impact on price but the observed market-adjusted price change on the day it was announced was sufficiently negative to nonetheless *pass* this test. Type II errors occur when the item of interest *did* in fact have an impact on price but the observed market-adjusted price change on the day it was announced *was not statistically significant*.

Again, these two types of errors are most easily understood by going back to our example. When the standard deviation is properly specified, as will be the case in normal economic times,¹⁸² the Type I error rate depends solely on the level of statistical confidence required by the test (for purposes of this discussion, 95%¹⁸³). To pass the test, the observed market-adjusted price change on the day of the item of interest must be at least -\$1.96. As depicted in Figure 9, there is only a 2.5% chance that the other firm-specific news and background noise had a net price impact this negative. Thus, there is only a 2.5% chance that we would observe a market-adjusted price change this negative if the item of interest in fact had no negative impact on price: The Type I error rate is 2.5%.¹⁸⁴ Observed market-adjusted negative price changes more negative than this—in our example, the observed change of -\$2.00 or -2.00%—pass the test.

182. See *infra* section III.B.1 (identifying magnitude of Type I and Type II errors in normal economic times).

183. Cf. Shari Seidman Diamond, Reference Guide on Survey Research, *in* Reference Manual on Scientific Evidence 359, 381 (3d. ed. 2011) [hereinafter *Manual on Scientific Evidence*] (“Traditionally, scientists adopt the 95% level of confidence”); David H. Kaye & David A. Freedman, Reference Guide on Statistics, *in* *Manual on Scientific Evidence*, *supra*, 211, 245 (noting 95% is confidence level “most popular” in scientific literature).

184. The standard methodology is to use a “two-tailed” test, which looks only at the magnitude of the price change. The result is considered statistically significant at the 95% level if the observed price change is greater than 1.96 times the standard deviation, whether positive or negative. The observed price change being statistically significant and negative implies that if the item of interest had no effect on price, there is no more than a 2.5% chance that we would observe a change this negative.

The Type II error rate—where the item of interest *did* in fact have a negative impact on price but the observed market-adjusted price change on the day it was announced *failed* the confidence test—is a bit more complicated. The Type II error rate depends not only on the level of statistical confidence required by the test, again 95% for this discussion, but also on the magnitude of the actual negative price impact of the item of interest and the issuer’s idiosyncratic volatility as measured by its standard deviation.

To see how this works, modify our continuing example slightly. Suppose, as depicted in Figure 10, that an item of interest in fact had a -1.00% market-adjusted impact on the issuer’s stock price and that, as before, the standard deviation of market-adjusted price changes for this issuer was 1.00% over the measuring period. Recall again that the market-adjusted price change observed on the disclosure day is the combination of the item of interest’s price impact and the net price impact of all the other firm-specific bits of news and background noise on that day.¹⁸⁵ The relevant question is: What is the likelihood that the observed market-adjusted price change on the disclosure day will be sufficiently negative to pass the 95% confidence standard, (i.e., more negative than -1.96%)? This will only happen if the net impact of all the other bits of firm-specific news and background noise on that day is at least -0.96% (which, when combined with the -1.00% impact of the item of interest, would in total be at least -1.96%).

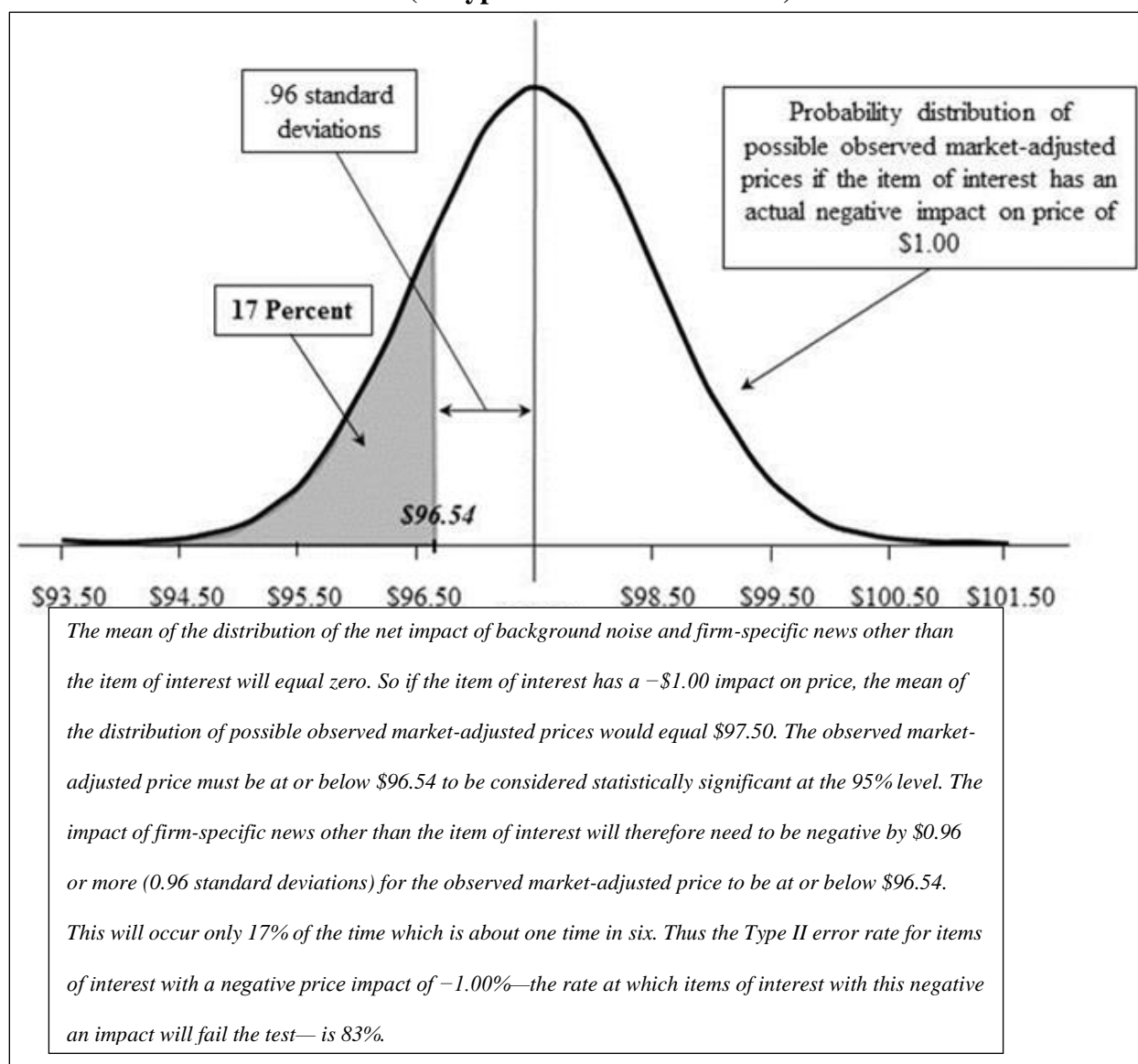
As depicted in Figure 10, this will be the case only about 17% of the time, or only about one time in six.¹⁸⁶ Thus, one cannot necessarily infer from an observed market-adjusted price change

185. The expected value of the effect on share price of these other factors is zero, but that only means that they are as likely to add to as to subtract from the negative effect on share price of the item of interest. As a highly simplified example, suppose there is one other piece of firm-specific information revealed to the market the same day. This other bit of news is unrelated to the news items of interest and if it were the only piece of firm-specific information revealed that day, would result in a positive price change of $+3\%$. The observed market-adjusted change in price will be 2% , even if, as posited, the news item of interest itself affected price by -1% .

186. For a more detailed discussion of the general approach to calculating the chances of this kind of error, see Edward G. Fox, Merritt B. Fox & Ronald J. Gilson, *Idiosyncratic Risk During Economic Downturns: Implications for the Use of Event Studies in Securities Litigation* 7–11 (Columbia Law & Econ., Working Paper No. 453, 2013)

failing the test that it likely did not affect price. Indeed, for an issuer with a standard deviation of 1%, an item of interest with an actual impact on price of -1.00% will fail the test most of the time.

Figure 10: Demonstration of Type II Error
An Item of Interest Having an Actual Negative Impact of \$1.00, the Observed Market Price Will Be Negative Enough to Be Considered Statistically Significant Only One Time in Six (A Type II Error Rate of 83%)



[hereinafter Fox, Fox & Gilson, Idiosyncratic Risk], <http://ssrn.com/abstract=2314058> [<http://perma.cc/7K5A-NGS3>] (expressing calculations mathematically).

More generally, the odds of an item of interest that actually had a negative effect on price passing a test based on any given level of confidence depends on the size of the actual effect relative to the standard deviation of past day-to-day market-adjusted price changes. Also, the higher the required confidence level, the lower the rate of false positives (Type I errors) and the higher the rate of false negatives (Type II errors). As we will see, each kind of error can have negative social consequences when event studies are used in securities litigation. That assessment leads to the matter of special concern here: Economic crisis-related idiosyncratic volatility spikes make the problem of Type I and Type II errors significantly worse.

1. The Magnitude of Type I and Type II Errors in Normal Times Using a 95% Confidence Level. — Our analysis starts by identifying the magnitude of Type I and Type II errors in normal times. This sets the baseline for our consideration of the impact of crisis-related volatility spikes.

a. Type I Errors. — In normal economic times, it is conventionally assumed that the same probability distribution generates the market-adjusted price changes on each of the approximately 250 trading days during the typical one-year observation period prior to the item of interest's disclosure.¹⁸⁷ Thus, the standard deviation of the observation period's approximately 250 market-adjusted price changes is a fairly precise estimate of the standard deviation of the probability distribution generating the net price impact of background noise and firm-specific information other than the item of interest on the day of its announcement. Under these circumstances, with a 95% confidence standard, the Type I error for a test of whether the item of interest had a negative impact on price is as calculated above: 2.5%.

187. A visual review of Figures 1 and 2, *supra*, provides an empirical verification of the reasonableness of the assumption. In these Figures, each moment's measure of idiosyncratic risk is the standard deviation of market-adjusted price changes over the previous twelve months. In other words it is a day-by-day look back over the prior twelve months. One can see that during normal times this measure stays quite steady from one day to the next.

b. Type II Errors. — Now consider what in normal times would be the Type II error rate associated with using the 95% confidence level if an item of interest in fact negatively affects the price by 5%.¹⁸⁸ Again, background noise and firm-specific information other than the item of interest will simultaneously affect the issuer’s market-adjusted price. Thus, the observed market-adjusted price change in any given instance may differ, one way or the other, from 5%. Type II error occurs when the observed market-adjusted price change—the combination of the price impact of the item of interest and these other impacts—is not sufficiently negative to pass the test of statistical significance. The starting point is to calculate how negative the market-adjusted price change accompanying the item of interest’s announcement needs to be to pass the test, just as was done in the example above concerning Type II error depicted in Figure 10. In Figure 6 above, market-cap-weighted average firm-specific volatility, as measured by variance, from the 1970s until the advent of the financial crisis, was in the range of 6% to 10% during noncrisis years, with an average of approximately 8%. This annualized variance translates to a daily standard deviation of 1.78%.¹⁸⁹ Again, for the observed market-adjusted price change to be considered statistically significant at the 95% level, the observed market-adjusted price change must be at least as negative as 1.96 times the standard deviation—it must be $-1.96 \times 1.78\% = -3.49\%$.

188. Securities lawyers often use 5% as a crude rule-of-thumb threshold percentage as a starting point for determining how much income or assets need to be misreported to be considered material. See *infra* note 243 and accompanying text (describing prevalence of 5% in securities litigation contexts).

189. Since our calculations are based on daily data, it is our annual variances that are interpolated using the following mathematical formula: $Var \sum_{i=1}^{252} \varepsilon_i = 252 * Var(\varepsilon_i)$ where $\sum_{i=1}^{252} \varepsilon_i$ is the sum of the market-adjusted returns on each of the 252 trading days each year, and thus the left-hand side is the annual variance of market-adjusted returns. The equality flows because the market-adjusted returns will be independent of one another in an efficient market. A reader can back out the daily variance by dividing the annualized numbers in Figure 2 by 252. The daily standard deviation is the square root of the daily variance. Due to the nonlinearity of variance, this is not the exact figure that we yield after marketcap weighting the standard deviations of the individual firms, but this difference is relatively minor.

The Type II error question is then how likely it is that the observed market-adjusted price be at least this negative when the actual price impact of the item of interest is -5% . The observed price change will be this negative unless the net impact of the other firm-specific news and background noise that day is $+1.51\%$ or more (so that, when added to the -5% impact of the item of interest, the combined impact is less negative than -3.49%). These other factors would have a net impact of $+1.51\%$ or more only about 20% of the time.¹⁹⁰ This is the Type II error rate—the likelihood that this item of interest, which actually negatively affected price by 5%, would not be accompanied by a market-adjusted price change sufficiently negative to be statistically significant. In these circumstances, a securities fraud claim will fail despite the fact that the disclosure did in fact have a 5% impact on stock price.¹⁹¹

Discussion can now move to the impact of crisis-induced increases in idiosyncratic volatility on event study methodology and hence on the conduct of securities litigation.

2. *Type I and II Errors in Periods of Crisis-Induced, High Idiosyncratic Risk.* — A spike in idiosyncratic volatility has two implications for Type I and Type II errors. First, if the spike

190. This calculation involves the distribution of possible observed values of the market-adjusted price change if the actual market-adjusted price impact of the tested item of news is -5% . The distribution of observed market-adjusted price changes accompanying tested items of news with an actual impact of -5% will approximate a normal distribution with a mean of -5% and a standard deviation equal to the standard error of 1.78%, representing the effect, plus or minus, of the other ordinary bits of firm-specific information that move the issuer's share price around every day. Since the observed change in prices will be considered statistically significant at the 95% level and have the right sign only if it is a decrease of greater than -3.49% , the question becomes: What are the chances that the observed change after imposition of the requirement will be at least this negative? The required negative change, -3.49% , is 0.85 standard deviations above -5% , and so, based on the normal distribution, there is then an 80.2% chance that the observed market-adjusted change in price will be a decrease greater than 3.49% and hence considered statistically significant at the 95% level in a two-tailed test. Thus, there is a 19.8% chance that the observed change will yield a false negative. In a related, more technical paper, we discuss in greater detail the general approach to calculating Type II errors. Fox, Fox & Gilson, *Idiosyncratic Risk*, supra note 186, at 7–11.

In fact, market-adjusted stock returns do not appear to converge exactly to a normal distribution (they are “fatter” tailed, i.e., they exhibit more extreme returns than if they followed a normal distribution). We use the normal distribution because it is used in virtually all event studies for testing statistical significance. The points outlined above and below are only strengthened if returns are “fat tailed” because, if returns are fat tailed for a given level of Type I error, one must set an even more stringent cutoff for statistical significance than one would under the normal distribution, which would induce more Type II errors.

191. A first impression is that the likelihood of a mistake is too high to provide the foundation for the litigation of securities fraud cases. While that assessment is beyond the scope of our effort here, it is worth keeping in mind the alternative: two expert financial analysts offering nonempirically grounded opinions where the error term is likely both greater and completely impossible to estimate with any precision.

recently occurred and conventional event study methodology is employed, the test of statistical significance that identifies the size of a price change large enough to meet the 95% confidence level can drastically underestimate the actual extent of Type I error. Put simply, the standard deviation used for determining the statistical significance of the price change will understate the real standard deviation because of the crisis-induced increase in idiosyncratic risk—that is, securities fraud claims will succeed when they would otherwise have failed. While more advanced techniques address this underestimation, they raise other problems.¹⁹² Second, more generally, even where the cutoff for statistical significance properly takes into account the crisis-induced spike in idiosyncratic risk and is set at the appropriate point to maintain the Type I error rate at 2.5%, a spike nonetheless can result in a several-fold increase in Type II error—that is, securities fraud claims will fail when they should have succeeded.

a. The Effect of a Recent Increase in Idiosyncratic Volatility. — Consider the situation where there has recently been a large increase in idiosyncratic volatility, as was the case in the fall of 2008 as the full dimensions of the financial crisis were just becoming apparent.¹⁹³ This special situation poses problems for the conventional strategy for testing whether an item of interest in fact affected price. In particular, it undermines the use of a one-year measuring period before the occurrence of the event in question as a proxy for the forces generating idiosyncratic volatility at the time the item of interest is announced.

The conventional event-study methodology assumes that the volatility in the company's market-adjusted stock price at the time of the announcement is the same as during the preceding

192. See *infra* notes 199–200 and accompanying text (explaining advanced Generalized Auto Regressive Conditional Heteroskedasticity (GARCH) model).

193. For example, during seven days in September 2008, Lehman Brothers failed and the federal government effectively nationalized AIG. Ronald J. Gilson & Reinier Kraakman, *Market Efficiency After the Financial Crisis: It's Still a Matter of Information Costs*, 100 Va. L. Rev. 313, 375 (2014) (noting full effect of capital market crisis had not been fully realized by dates of Lehman Brothers failure and federal government's intervention in AIG).

one-year observation period.¹⁹⁴ In the special situation of a recent crisis-induced increase in idiosyncratic volatility, this assumption fails radically, resulting in an insufficiently negative cutoff and securities fraud actions succeeding when they should not.

b. Potential Understatement of Type I Errors. — In this special situation, the standard deviation of market-adjusted price changes during the one-year observation period will underestimate the daily net price fluctuation by the time of the item of interest's announcement. This is because of the sharp increase in idiosyncratic risk that accompanies each crisis.¹⁹⁵ With the cutoff for what is considered statistically significant then incorrectly based on a standard deviation smaller than what actually prevailed at the time of the event, more than 2.5% of the tested news items that in fact had no effect on price will be accompanied by market-adjusted price changes that satisfy this erroneous cutoff. Thus, the Type I error rate will be understated to the benefit of plaintiffs in securities fraud actions.

A hypothetical provides a sense of the extent of the understatement of Type I errors. Assume that the standard deviation of firm-specific price changes during the observation period was 1.5% (the standard deviation for the average S&P 100 firm for the July 1, 2007, to June 30, 2008, period). Assume as well that the idiosyncratic volatility had increased to 2.7% (the standard deviation for such a firm in the July 1, 2008, to June 30, 2009, period) and that the event of interest occurs early in this second period. The observation period is thus July 1, 2007, to June 30, 2008. Using the conventional methodology (a one-year pre-event estimation period), any observed market-adjusted price drop of 2.95% ($1.96 \times 1.5\%$) or more on the date the item of interest is disclosed would be statistically significant at the 95% confidence level. But by the date

194. See *supra* section III.A.2 (describing conventional method of using standard deviation of observation period's approximately 250 market-adjusted price changes as measure of the probability distribution generating the net price impact of background noise and firm-specific information other than the item of interest).

195. See *supra* Part I (summarizing empirical findings).

of disclosure the standard deviation of idiosyncratic volatility has risen to 2.7%. The result is that about 14% of the time, items of interest with no effect on price will be accompanied by market-adjusted price changes sufficiently negative to be statistically significant.¹⁹⁶ In effect, our event study is comparing apples and oranges: the current higher volatility with a cutoff based on the older lower volatility level. *Thus, the Type I error rate would be almost six times the supposed 2.5% level.* As will be discussed in section IV.B.2, liability clearly should not be imposed on an issuer in a securities fraud claim where the misstatement in fact had no impact on price. Yet, with a recent increase in idiosyncratic volatility of this sort, using a conventionally designed event study to determine loss causation would increase six-fold the likelihood of liability being incorrectly imposed in such a situation.

c. Correcting for Understated Type I Errors and the Resulting Increase in Type II Errors. —

One approach to addressing the problem of increased Type I error caused by the increase in idiosyncratic volatility after the estimation period ends is to begin the estimation period only after the idiosyncratic risk increases. Where the tested item of news occurs relatively soon after the increase in volatility, however, an estimation period beginning with the volatility increase and ending immediately before the item of interest will be shorter, possibly much shorter, than the conventional one-year period. This will result in less precise estimates of the relationship between the issuer's stock price and the ups and downs of the market as a whole (i.e., the issuer's β)¹⁹⁷ and of the level of idiosyncratic risk. These less precise estimates ameliorate Type I error, but only by increasing Type II error.¹⁹⁸

196. This percentage reflects that -2.95% is 1.09 times the actual standard deviation of 2.70% . Based on standard statistical tables for the normal distribution, 14% of all outcomes are more negative than 1.09 standard deviations below the mean.

197. See Appendix, *supra* note 144, at 11 (showing how this will increase idiosyncratic risk).

198. One way to counter this problem of an abbreviated observation period is to extend the period forward beyond the date the tested item of news becomes public, but omit from the sample the price change on that date to

Another approach to addressing the conventional strategy's significant understatement of Type I error when there is a recent increase in idiosyncratic volatility is to use a full one-year observation period that ends before the tested item of news becomes public, but to use more advanced econometrics that take into account the midstream change in idiosyncratic volatility.¹⁹⁹ The additional complexity of these techniques makes them harder for a court, let alone a jury, to evaluate. In particular, these models offer the expert performing the analysis relatively wider scope for important choices—that is, choices that importantly affect whether the securities fraud claim will succeed.²⁰⁰ In such circumstances, it is predictable that the plaintiff's and defendant's experts will disagree. The court may be ill-equipped to determine which expert to believe, and a jury will likely be utterly confused by what is a very sophisticated econometric debate.

d. The Effect of an Increase in Idiosyncratic Volatility More Generally. — If disclosure of the item of interest occurs at a time when idiosyncratic risk is sharply elevated by economic crisis, the Type II error rate for any maximum acceptable rate of Type I error will be much higher than it would be in normal times. While this problem begins as soon as there is an increase in idiosyncratic risk, for expository simplicity, we will disentangle it from the problem just discussed above by assuming that idiosyncratic risk has remained stable for more than a year

avoid having the event whose effect the study is trying to determine influence the measure. The downside is that the volatility induced by the tested item of news may seep into other days in the observation period, thus again increasing Type II error.

199. One method for dealing with shifting volatility is to use a GARCH model, which makes current volatility a function of past volatility and net-of-market returns. See Appendix, *supra* note 144, at fig.(2)A-9 (providing more detailed explanation and analysis using GARCH). GARCH is used frequently in academic studies of stock price volatility. See, e.g., A. Corhay & A. Tourani Rad, Conditional Heteroskedasticity Adjusted Market Model and an Event Study, 36 Q. Rev. Econ. & Fin. 529, 530 (1996) (applying GARCH in event study of divestitures). However, its use in litigation remains relatively rare. Other possibly useful methods flow from the literature identifying structural breaks in the data. See generally, e.g., Jushan Bai & Pierre Perron, Estimating and Testing Linear Models with Multiple Structural Breaks, 66 *Econometrica* 47 (1998).

200. For example, the expert must choose how many autoregressive terms to include in GARCH or how to identify the structural break. See Corhay & Rad, *supra* note 199, at 531–32 (providing model for estimating structural break).

before the item of interest occurs. This treatment controls for the mismatch between volatility levels during the estimation period and at the time the event occurs.

Higher idiosyncratic volatility during the estimation period results in a larger standard deviation of the issuer's past market-adjusted share price changes and consequently, a larger standard deviation with which to evaluate the price effect of the item of interest's disclosure. As a result, the cutoff for an observed price change to be considered statistically significant must be more negative. Thus, for any given maximum acceptable level of Type I error, the *observed* change in market-adjusted price accompanying the particular event must be more negative for the change to be sufficiently negative to meet the required level of statistical confidence.²⁰¹ Consequently, the chance that an *actual* negative price effect of any given size will be accompanied by an observed price change that meets this standard is reduced. In other words, the greater the volatility, the greater the number of Type II errors for any given allowable number of Type I errors.²⁰² In essence, the same level of price effect is harder to detect when more is going on in the background.²⁰³

The spike in idiosyncratic risk accompanying the recent financial crisis provides a dramatic illustration of this point. Recall the hypothetical discussed above that considered the level of Type II errors in *normal* times when an event in fact has a -5% impact on price.²⁰⁴ Now consider an event with the same actual price impact of -5% but that occurs during *crisis* times with a

201. For a formal demonstration of this point, see Fox, Fox & Gilson, *Idiosyncratic Risk*, *supra* note 186, at 7–11.

202. *Id.* at 11 (showing mathematically that “if the Type I error rate is maintained at 2.5%, the Type II error rate jumps in crisis times from 19.8% to 66%”).

203. See *id.* (“[I]n high volatility times resembling the 2008–09 financial crisis, only a bit more than one in three disclosures that actually affected an issuer’s price by 5% would pass the test of being considered statistically significant at the 95% level, compared with . . . [over] four out of five . . . in normal times.”).

204. See *supra* section II.B.1.b (presenting hypothetical scenario).

spike in idiosyncratic risk of the magnitude observed during the recent financial crisis.²⁰⁵ Then the market-adjusted price change needed to meet the 95% confidence level standard jumps from -3.49% to -6.33%. This implies an increase in Type II error rate from 20% to 66%.²⁰⁶ So, in such high volatility times, about only one in three items of interest whose disclosure in fact affects an issuer's share price by -5% would be accompanied by observed price changes considered statistically significant at the 95% level, compared with four out of five passing the test in normal times. More generally, whatever the level of maximum allowable Type I errors and whatever the actual impact of a misstatement on price, higher volatility results in a higher level of Type II errors.²⁰⁷

IV. LEGAL IMPLICATIONS OF CRISIS-INDUCED SPIKES IN IDIOSYNCRATIC RISK

Crisis-induced spikes in idiosyncratic risk can have important implications for corporate and securities law both through their effect on event studies and more generally. In this Part, we examine three examples: In section IV.A, establishing causation and materiality in fraud-on-the-market securities class actions; in section IV.B, establishing materiality in other securities law

205. For this example, we assume that the standard deviation increases from the normal-times level of 1.78%, see *supra* note 189, to 3.23%, approximately equal to the daily idiosyncratic volatility of the market cap-weighted average firm in the S&P 100 during the peak of the financial crisis.

206. Similarly to the calculations in *supra* note 190, this calculation involves the distribution of possible observed values of the market-adjusted price changes if the actual market-adjusted price impact of the corrective disclosure is -5%. The distribution of observed market-adjusted price changes accompanying corrective disclosures with an actual impact of -5% will approximate a normal distribution with a mean of -5% and a standard deviation equal to the standard deviation of 3.23%, representing the effect, plus or minus, of the ordinary bits of firm-specific information and background noise that affects the issuer's share price every day during this period of high volatility. Since the observed change in prices will be considered statistically significant at the 95% level and have the right sign only if it is a decrease of 6.33%, the question becomes: What are the chances that the observed change will be of that magnitude? The required negative change, -6.33%, is 0.41 standard deviations beyond -5%. Based on standard statistical tables for the normal distribution, there is then only a 34% chance that the observed change in market-adjusted prices will be a decrease of at least 6.33% and hence statistically significant at the 95% level. Thus, there is a 66% chance that the observed change will yield a false negative. For a more detailed discussion of these points, see Fox, Fox & Gilson, *Idiosyncratic Risk*, *supra* note 186, at 10-11.

207. Fox, Fox & Gilson, *Idiosyncratic Risk*, *supra* note 186, at 7-9 ("The greater SD [standard deviation], for any given [actual impact on price], the greater the likelihood of Type II error.").

contexts; and in section IV.C, judicial deference to board rejections of premium acquisition offers and the substantive coercion defense.

A. Fraud-on-the-Market Shareholder Class Actions: Establishing Causation and Materiality

The spike in idiosyncratic risk that occurs during financial crises causes a very substantial decline in the usefulness of fraud-on-the-market class actions in crisis times.²⁰⁸ These actions allow buyers in secondary securities markets to recover from the issuer losses that they incurred by purchasing at prices inflated by the issuer’s misstatements, without individual class members having to prove that they actually relied upon (or even knew about) the misstatement giving rise to their claim.²⁰⁹ These fraud-on-the-market actions, based on alleged violations of section 10(b) of the Securities and Exchange Act of 1934 (Exchange Act) and Rule 10b-5 promulgated thereunder, currently produce the bulk of all the damages paid out in settlements and judgments pursuant to private litigation under the U.S. securities laws.²¹⁰ Plaintiffs in these fraud-on-the-market actions depend on the reliability of event studies. Idiosyncratic-risk spikes diminish that reliability for studies conducted during economic-crisis periods.

208. A fraud-on-the-market action is an implied right of action for civil damages based on a misstatement made in violation of the Securities Exchange Act of 1934, Pub. L. No. 73-291, 48 Stat. 881 (codified as amended at 15 U.S.C. §§ 78a--78pp (2012)), 15 U.S.C. § 10(b) and Rule 10b-5, 17 C.F.R. § 240.10b-5 (2015), and 15 U.S.C. § 78j(b), promulgated thereunder. Rule 10b-5 provides in relevant part:

It shall be unlawful for any person, directly or indirectly, . . . [t]o make any untrue statement of a material fact or to omit to state a material fact necessary in order to make the statements made, in the light of the circumstances under which they were made, not misleading . . . in connection with the purchase or sale of any security.

17 C.F.R. § 240.10b-5. Any statement made by an issuer that is “reasonably calculated to influence the investing public,” for example by being made to the media, satisfies Rule 10b-5’s requirement that it be “in connection with the purchase or sale of a security,” even though neither the issuer nor its officials buy or sell shares themselves. SEC v. Tex. Gulf Sulphur Co., 401 F.2d 833, 862 (2d Cir. 1968) (en banc).

209. See *Basic Inc. v. Levinson*, 485 U.S. 224, 241–42, 250 (1988) (holding lower courts may apply fraud-on-the-market presumption on rebuttable basis).

210. See Fox, *Securities Class Actions*, supra note 133, at 1176 & n.2 (citing interviews with practitioners and quantitative data about initial complaints).

The centrality of securities fraud class actions dates to the Supreme Court’s decision in *Basic Inc. v. Levinson*.²¹¹ Prior to that decision, courts required each plaintiff to prove that a misrepresentation was “a substantial factor in determining the *course of conduct* which results in [the recipient’s] loss.”²¹² Under this traditional, pre-*Basic* rule, securities fraud class actions were extremely difficult to prosecute. Absent the aggregation of claims (and associated economies of scale in litigation costs) that a class action allows, pursuing a securities fraud claim is infeasible for all but the largest traders. But certification of a shareholder class action seeking money damages requires that common issues of fact and law predominate,²¹³ a requirement that cannot be met if each plaintiff must individually prove reliance and causation.²¹⁴

Basic fundamentally changed the manner in which causation could be proved. Under its then-new “fraud-on-the-market” theory, a material misstatement by an issuer whose shares trade in an efficient market is expected to affect the issuer’s share price.²¹⁵ The Court said that because such misrepresentations will have an impact on the security’s price and because all traders rely on the price, individual reliance can be presumed rather than proven, thereby eliminating the need for proof of individual reliance.²¹⁶ This presumption makes class actions economically

211. 485 U.S. 224.

212. *List v. Fashion Park, Inc.*, 340 F.2d 457, 462 (2d. Cir. 1965) (emphasis added) (internal quotation marks omitted) (quoting Restatement of Torts § 546 (Am. Law Inst. 1938)).

213. Fed. R. Civ. P. 23(b)(3).

214. See *Castano v. Am. Tobacco Co.*, 84 F.3d 734, 745 (5th Cir. 1996) (“[A] fraud class action cannot be certified when individual reliance will be an issue.”).

215. See *Basic*, 485 U.S. at 246–47 (endorsing presumption that “market price of shares traded on well-developed markets reflects all publicly available information, and, hence, any material misrepresentations”).

216. *Id.* at 247 (“Because most publicly available information is reflected in market price, an investor’s reliance on any public material misrepresentations, therefore, may be presumed for purposes of a Rule 10b-5 action.”).

feasible, with the concomitant large potential exposure to issuers. The result was an enormous growth in securities fraud class action litigation.²¹⁷

This is where event studies come in. They are the predominant way that class action plaintiffs establish both the materiality of the misstatement and the causal link between that misstatement and their losses—that the issuer’s misstatement in fact inflated the prices the plaintiffs paid for their shares.²¹⁸ Absent crisis-induced increases in idiosyncratic volatility, this methodology works reasonably well. But our finding of very large spikes in idiosyncratic volatility at times of economic crisis shows that it works less well during these periods, precisely the time when large price movements in individual stocks are likely to give rise to an increase in securities fraud actions. This reduced effectiveness of event studies during such periods raises important issues, discussed below, concerning both how courts should administer fraud-on-the-market suits in such times and how these suits should fit into the larger framework of securities law enforcement.

1. Fraud-on-the-Market Actions and the Importance of the Misstatement’s Effect on Price.

— The first step in understanding the issues raised by our empirical results for fraud-on-the-market actions is to examine how such an action works.

217. In the five-year period beginning April 1988, shortly after the Supreme Court’s decision in *Basic*, and ending March 1993, the total cash amount paid to settle federal class actions alleging that issuer misstatements distorted share price was \$2.5 billion. Vincent E. O’Brien & Richard W. Hodges, A Study of Class Action Securities Fraud Cases 1988–1993 I-5 (1993) (unpublished study) (on file with the *Columbia Law Review*). By the early 2000s, the total amount of such settlements had increased dramatically. Indeed, the value of settlements paid from January 2005 through December 2007 (a period that includes the three blockbuster cases of WorldCom, Enron, and Tyco), including the disclosed value of any noncash components, totaled over \$39.5 billion. Laarni T. Bulan, Ellen M. Ryan & Laura E. Simmons, Cornerstone Research, Securities Class Action Settlements: 2014 Review and Analysis 3 fig.2 (2015), <http://securities.stanford.edu/research-reports/1996-2014/Settlements-Through-12-2014.pdf> [http://perma.cc/UAH6-SJPM]. This figure decreased to \$9.2 billion for the three years beginning January 2012 and ending December 2014. Id.

218. A recent article sympathetic to plaintiffs explains the conclusion that an event study is mandatory for a securities class action case to proceed. See Michael J. Kaufman & John M. Wunderlich, Regressing: The Troubling Dispositive Role of Event Studies in Fraud Litigation, 15 *Stan. J.L., Bus. & Fin.* 183, 187 (2009) (“The interrelated questions of materiality, reliance, loss causation, and damages all require an event study for their resolution. The overriding substantive issue in securities fraud cases has become whether an expert has proffered an opinion based on a reliable event study.”).

a. The Stripped-Down Model of Fraud-on-the-Market Litigation. —Analyzed doctrinally, the legal issues in a fraud-on-the-market securities action (assuming the plaintiff establishes at the outset a public misstatement made with scienter by an issuer whose shares trade in an efficient market) are the materiality of the misstatement, loss causation, transaction causation, and damages.²¹⁹ The litigation of such a claim can be described in terms of the allocation, between the parties, of the burdens of proof and persuasion on each of these issues at each stage of the litigation and the allowable forms of evidence.²²⁰ For our purposes here, however, these issues reduce to two: Did the misstatement inflate the price paid by more than a de minimis amount, and if so, did the plaintiff suffer a loss as a result?

Where both these questions can be answered affirmatively, all the doctrinal elements for the cause of action will be satisfied.²²¹ An investor who purchases shares of the issuer while the price is inflated by the misstatement, and who still holds the shares at the time the truth is revealed, has suffered a loss. She paid too much as a result of the misstatement, and because revelation of the truth dissipates this inflation, she has not been able to recoup her loss by selling into a still-inflated market. Her loss thus satisfies the loss causation requirement (with transaction causation

219. See, e.g., *Dura Pharm., Inc. v. Broudo*, 544 U.S. 336, 341–42 (2005) (“[T]he action’s basic elements include: (1) a material misrepresentation (or omission); (2) scienter . . . ; (3) a connection with the purchase or sale of a security; (4) reliance [or] ‘transaction causation’; (5) economic loss; and (6) ‘loss causation’” (emphasis omitted) (internal citations omitted)).

220. See *id.* at 342–43 (noting plaintiff must independently establish causation and “an inflated purchase price will not itself constitute or proximately cause the relevant economic loss” because of “tangle of factors affecting price”).

221. This observation parallels Daniel Fischel’s insight in a seminal pre-*Basic* article that commented on lower court cases that were the origin of the fraud-on-the-market cause of action. Daniel R. Fischel, *Use of Modern Finance Theory in Securities Fraud Cases Involving Actively Traded Securities*, 38 *Bus. Law.* 1, 12–13 (1982). Fischel suggested that the adoption of the cause of action reflected an underlying view of the market that most investors were price takers. As a consequence, they are hurt by a misstatement because of its effect on price, not its effect on their decisions to buy or sell. Fischel observed that for an action based on this view, the traditional doctrinal issues of materiality, reliance, and damages reduce to a single inquiry: Did the misstatement affect price and if so by how much? *Id.* at 13. The Supreme Court cited Fischel’s article in *Basic*. 485 U.S. at 246 n.24.

For an example of a judicial opinion explicitly endorsing this collapsing of the three traditionally separate doctrinal elements into a single empirical test, see *In re Verifone Sec. Litig.*, 784 F. Supp. 1471, 1479 (N.D. Cal. 1992) (noting fraud-on-the-market theory “subsumes” reliance, materiality, causation, and damages inquiries into single analysis), *aff’d*, 11 F.3d 865 (9th Cir. 1993).

being satisfied as well by the mere fact that the situation receives the fraud-on-the-market presumption).²²² The amount of this loss is her damages, thereby satisfying the damages requirement.²²³ As for the materiality requirement, the Supreme Court has held that a fact is material if there is a substantial likelihood that a reasonable investor would consider it important in a decision whether to purchase or sell a security.²²⁴ Anytime a misstatement meaningfully inflates the price of a security trading in an efficient market, it has obviously had an actual effect on the behavior of investors. This strongly suggests that a reasonable investor, like those actually trading in the market, would have found it important.²²⁵

b. Focus on the Corrective Disclosure. — Where an issuer’s original misstatement hides a truth that is less favorable than the market’s expectations for the issuer at the time the

222. In the pre-fraud-on-the-market years before *Basic*, the courts refined their causation analysis to require two showings: transaction causation and loss causation. Transaction causation required the plaintiff to show she would not have purchased but for the misstatement. Loss causation required the plaintiff to show that the untruth was responsible for the loss in some reasonably direct or proximate way. See Merritt B. Fox, After *Dura*: Causation in Fraud-on-the-Market Actions, 31 J. Corp. L. 829, 834–36 (2006) [hereinafter Fox, After *Dura*] (discussing pre-*Basic* framework). These concepts do not fit well with the alternative causal connection allowed in the fraud-on-the-market actions, but the courts have maintained the two requirements. Transaction causation is presumed in any situation where the fraud-on-the-market presumption is allowed: where there is a material misstatement by an issuer whose shares trade in an efficient market. See, e.g., *Semerenko v. Cendant Corp.*, 223 F.3d 165, 178–83 (3d Cir. 2000) (“[A] plaintiff in a securities action is generally entitled to a rebuttable presumption of reliance if he or she purchased or sold securities in an efficient market.”). A showing of loss causation requires not only that the misstatement inflated the issuer’s share price but also that there was a causal connection between this inflation and a loss by the plaintiff. *Dura Pharm.*, 544 U.S. at 346–48 (2005) (holding plaintiffs did not adequately allege “proximate causation and economic loss”). Thus, the basic causal inquiry in the fraud-on-the-market theory is framed in terms of loss causation.

223. This corresponds to the “out-of-pocket” measure of damages that is standard in Rule 10b-5 cases. See *Green v. Occidental Petroleum Corp.*, 541 F.2d 1335, 1341–46 (9th Cir. 1976) (Sneed, J., concurring in part and concurring in the result in part) (describing out-of-pocket measure of damages and its application).

224. *Basic*, 485 U.S. at 232 (“We now expressly adopt the *TSC Industries* standard of materiality for the § 10(b) and Rule 10b-5 context.”). In *TSC Industries, Inc. v. Northway, Inc.*, the Supreme Court found that a fact is material “if there is a substantial likelihood that a reasonable shareholder would consider it important in deciding how to vote.” 426 U.S. 438, 449 (1970). Materiality “does not require proof of a substantial likelihood that disclosure of the omitted fact would have caused the reasonable investor to change his vote.” *Id.* Rather, it contemplates “a showing of a substantial likelihood that, under all the circumstances, the omitted fact would have assumed actual significance in the deliberations of the reasonable shareholder.” *Id.* The Court went on to say, “Put another way, there must be a substantial likelihood that the disclosure of the omitted fact would have been viewed by the reasonable investor as having significantly altered the ‘total mix’ of information available.” *Id.*

225. See, e.g., *Teamsters Loc. 445 Freight Div. Pension Fund v. Bombardier Inc.*, 546 F.3d 196, 207–08 (2d Cir. 2008) (noting “it is difficult to presume that the market will integrate the release of material information about a security into its price” absent evidence of efficient market for that security); *In re Merck & Co. Sec. Litig.*, 432 F.3d 261, 273–74 (3d Cir. 2005) (noting “‘reasonable investors’ are the market” and in efficient market, information is reflected in price, so when information changes price, it must be important to reasonable investors).

misstatement is made, the misstatement will not increase the issuer's share price. Instead, it just avoids the share-price decline that would have resulted from the truth. Because a substantial portion of all alleged issuer misstatements are of this kind,²²⁶ the inquiry into materiality and causation usually focuses on the market's reaction when the truth (the corrective disclosure) comes out. If the announcement of the truth causes the price to decline, the logic goes, the misstatement must have previously inflated the issuer's share price relative to what it would have been had the misstatement not been made. And in an efficient market, the disclosure of the truth guarantees that the inflation in the stock's price has been dissipated.²²⁷

2. *The Use of an Event Study to Establish that a Corrective Disclosure Has Had a Negative Impact on Price.* — Event studies are commonly used in fraud-on-the-market suits for assessing the likelihood that a corrective disclosure has in fact negatively affected price by more than a de minimis amount.²²⁸

a. *Event Studies as Evidence of Loss Causation and Materiality.* — For investors who purchase an issuer's stock after the issuer's misstatement and still hold the stock at the time of the corrective disclosure, an event study that reveals a statistically significant, market-adjusted

226. See, e.g., *SEC v. Todd*, 642 F.3d 1207, 1220–21 (9th Cir. 2011) (finding possible “material misrepresentations” where issuer’s CEO “failed to publicly disclose in an analysts’ conference call and an earnings press report that [issuer] Gateway’s ability to meet analysts’ revenue growth expectations was based largely on . . . one-time . . . transactions”); *Chamberlain v. Reddy Ice Holdings, Inc.*, 757 F. Supp. 2d 683, 702–04 (E.D. Mich. 2010) (noting plaintiff-shareholders pled materially false and misleading statements with sufficient particularity by alleging issuer hid violation of antitrust laws).

227. See Brealey, Myers & Allen, *supra* note 123, at 324–25 (summarizing “semistrong market efficiency” hypothesis, according to which “prices reflect . . . all . . . public information”).

228. See, e.g., *In re Sadia, S.A. Sec. Litig.*, 95 Fed. Sec. L. Rep. 806 (CCH) (S.D.N.Y. 2010) (“Vellrath conducted an event study of the daily movements in the price of Sadia’s ADRs [before and after issuer’s fraud-related corrective disclosure] with the goal of identifying any ‘abnormal returns’, i.e., returns greater than or less than . . . one would expect on the security . . .” (quoting Expert Report of Marc Vellrath, Ph.D., CFA, ¶41, *In re Sadia, S.A. Sec. Litig.*, 95 Fed. Sec. L. Rep. at 806 (No. 1:08-CV-09528 (SAS)), 2009 WL 5164437)); *Gen. Elec. Co. v. Jackson*, 595 F. Supp. 2d 8, 22–24 (D.D.C. 2009) (summarizing litigants’ use of event studies); *In re Seagate Tech. II Sec. Litig.*, 843 F. Supp. 1341, 1368 (N.D. Cal. 1994) (“Defendants’ expert . . . conducted the econometric ‘event study’ by noting the movement of the price of [issuer] Seagate common for each day in the class period and comparing it with an industry index.”); see also Kaufman & Wunderlich, *supra* note 218, at 187 (arguing “properly conducted event study . . . has become a substantive and essential element of a securities fraud claim itself”).

decline in price when the truth is revealed provides evidence of both materiality and loss causation.²²⁹ Indeed, some courts explicitly require the plaintiff to present, through expert testimony, an event study in order to make a showing of loss causation.²³⁰

Conforming to the usual social science convention,²³¹ courts generally have adopted the 95% confidence level as the standard to be used in securities litigation for determining the price effect of a corrective disclosure.²³² In choosing this standard, the courts are in essence establishing the

229. See, e.g., *In re REMEC Inc. Sec. Litig.*, 702 F. Supp. 2d 1202, 1266, 1275 (S.D. Cal. 2010) (noting “plaintiff must show . . . economic loss occurred after the truth behind the misrepresentation or omission became known to the market” and “decline in stock price caused by the revelation . . . must be statistically significant”); *In re Imperial Credit Indus., Inc. Sec. Litig.*, 252 F. Supp. 2d 1005, 1015–16 (C.D. Cal. 2003) (granting defendant issuer’s motion for summary judgment because plaintiff failed to provide “event study or similar analysis . . . [to] eliminate that portion of the price decline . . . which is unrelated to the alleged wrong” and is instead attributable to “market events for which Defendants cannot be held responsible”), *aff’d sub nom. Mortensen v. Snively*, 145 F. App’x 218 (9th Cir. 2005). Simply showing that the misstatement inflated price is not sufficient to establish loss causation. *Dura Pharm., Inc. v. Broudo*, 544 U.S. 336, 342–46 (2005) (holding plaintiffs must prove proximate cause and economic loss caused by misrepresentation).

230. See, e.g., *Fener v. Operating Eng’rs Const. Indus. & Misc. Pension Fund (LOCAL 66)*, 579 F.3d 401, 409 (5th Cir. 2009) (“Although analyst reports and stock prices are helpful in any inquiry, the testimony of an expert—along with some kind of analytical research or event study—is required to show loss causation.”). Plaintiffs must generally provide this evidence at the summary judgment stage. See Mary K. Warren & Sterling P.A. Darling, Jr., *The Expanding Role of Event Studies in Federal Securities Litigation*, 6 No. 6 Sec. Litig. Rep. 19 (2009) (“[A]n event study is commonly the device that creates a triable question of fact The presence or absence of an event study may, therefore, result in summary judgment for the defendant or, in limited circumstances, the plaintiff.”). With respect to materiality, see *In re Gaming Lottery Sec. Litig.*, Fed. Sec. L. Rep. (CCH) ¶90,763, at 93,716, 93,716 (S.D.N.Y. Feb. 16, 2000) (event studies are “accepted method for the evaluation of materiality”); William O. Fisher, *Does the Efficient Market Theory Help Us Do Justice in a Time of Madness?*, 54 *Emory L.J.* 843, 871, 874–83 (2005) (collecting cases).

231. See Michael Cowles & Caroline Davis, *On the Origins of the 0.05 Level of Statistical Significance*, 37 *Am. Psychol.* 553, 553 (1982) (noting “conventional use of the 5% level as the maximum acceptable probability for determining statistical significance” in the social sciences and exploring convention’s origin); Kaye & Freedman, *supra* note 183, at 251 (“In practice, statistical analysts typically use levels of 5% and 1%. The 5% level is most common in social science, and an analyst who speaks of significant results without specifying the threshold probably is using this figure.”).

232. See, e.g., *United States v. Hatfield*, 795 F. Supp. 2d 219, 234 (E.D.N.Y. 2011) (recognizing, in criminal securities fraud case, “95% confidence interval is the threshold typically used by academic economists in their work”); *In re REMEC Inc. Sec. Litig.*, 702 F. Supp. 2d at 1266 (to establish loss causation, event study must show “decline in stock price caused by the revelation of that truth [is] statistically significant” (citations omitted)); Cornerstone Research, *Estimating Recoverable Damages in Rule 10b-5 Securities Class Actions 9* (2014), <https://www.cornerstone.com/GetAttachment/df483ce3-5fe3-41b8-ad83-f704dc8426f7/Estimating-Recoverable-Damages-in-Rule-10b-5-Securities-Class-Actions.pdf> [http://perma.cc/8ZR3-EYNK] (last visited Oct. 23, 2015) (“A confidence interval of 95 percent is often applied in academic event studies and frequently accepted by courts.”). Kaye & Freedman, *supra* note 183, at 251 n.101, note that the Supreme Court “implicitly” endorsed the 95% confidence level in two 1977 decisions. See also *Hazelwood Sch. Dist. v. United States*, 433 U.S. 299, 311 n.17 (1977) (finding shortfall of black teachers hired by defendant school district “significant” because it represented “difference of more than three standard deviations”); *Castaneda v. Partida*, 430 U.S. 482, 496 & n.17 (1977) (finding sufficient “proof . . . to establish a prima facie case of discrimination against . . . Mexican-Americans in . . . grand jury selection” where statistical evidence showed the observed number of Mexican American jurors was less than expected number by “greater than two or three standard deviations”). Kaye and Freedman point out that “[a]lthough the Court did not say so,” the differences treated as “significant” in

plaintiff's burden of persuasion concerning whether the corrective disclosure in fact affected the issuer's share price. The "null hypothesis" is that the corrective disclosure had no negative influence on price. If the plaintiff can persuade the fact-finder that the null hypothesis can be rejected with at least 95% confidence, the plaintiff has established both loss causation and materiality.

b. The Working Assumption that the 95% Standard Maximizes Social Welfare in Normal Times. — We will assume here that this adoption of the 95% standard in fraud-on-the-market suits is the correct, social-welfare-maximizing standard, at least in normal times.²³³ For our purposes, the assumption that 95% is the correct standard in normal times forms a useful baseline that conforms to current practice.²³⁴ This, in turn, allows a focus on the ultimate concern here: whether an economic crisis-induced spike in idiosyncratic risk changes the appropriate confidence-interval standard for event studies, regardless of what standard might be optimal in normal times.

i. The Choice of Confidence Level Sets the Terms of the Tradeoff Between Type I and Type II Error. — An event study involves inevitable tradeoffs between Type I error and Type II error, with the terms of this tradeoff determined by the test's chosen level of statistical confidence.²³⁵ Recall that if the 95% confidence level is chosen, an event study will have a 2.5% Type I error rate, i.e., there is a 2.5% chance that a corrective disclosure that in fact has no effect on price will

Castaneda and Hazelwood School District "produce p-values of about 5% and 0.3% when the statistic is normally distributed." Kaye & Freedman, *supra* note 183, at 251 n.101.

233. Although beyond the scope of this Article, this assumption is itself worthy of further consideration. See, e.g., Leslie A. Demers, *Lost in Transposition: Erroneous Conflation of Statistical Certainty with Evidentiary Standards* 3 (Feb. 27, 2013) (unpublished manuscript), <http://ssrn.com/abstract=2225998> (on file with the *Columbia Law Review*) (describing errors courts make with 95% confidence-interval standard).

234. See *supra* note 232 (listing cases that applied 95% confidence-interval standard).

235. See *supra* section III.B (defining Type I and Type II errors and explaining tradeoffs between them).

be accompanied by an observed market-adjusted price change negative enough to pass the significance test (and allow the suit to go forward).

The choice of the 95% confidence level also influences the Type II error rate, i.e., the chance that a corrective disclosure that in fact has a negative impact on price will be accompanied by an observed market-adjusted price that is insufficiently negative to meet the required confidence level (and so will not allow a class action suit to go forward). But an event study's Type II error rate depends as well on how negative the disclosure's actual price impact is and on the level of idiosyncratic volatility associated with the issuer's stock. As we have seen, a corrective disclosure with a -5% price impact made by an issuer with what, in normal times, is the average level of idiosyncratic volatility, the choice of the 95% confidence level implies a Type II error rate of about 20%.

More generally, holding constant the level of idiosyncratic volatility, the choice of the 95% confidence level implies a Type I error rate of 2.5% and a set of Type II error rates that correspond to the magnitude of a corrective disclosure's actual negative impact on price. The more negative the actual impact, the lower the Type II error. In comparison, a confidence level stricter than 95% would imply a lower Type I error rate and a set of Type II error rates that, for each possible magnitude of negative impact on price, would be higher than it would be with the choice of the 95% confidence level. A confidence level less strict than the 95% level would have the opposite effects.

Thus, holding constant the level of idiosyncratic volatility, there is, for any given negative impact on price, a tradeoff between Type I and Type II error rates, with the point on this tradeoff determined by how strict the chosen confidence level is. Thus, the choice of the confidence level is the choice of a set of tradeoff points, each corresponding to a particular magnitude of a

corrective disclosure's actual negative impact on price. If idiosyncratic risk increases, the terms of this tradeoff worsen. For a given chosen level of confidence and Type I error rate, the Type II error rate corresponding to any particular magnitude of a corrective disclosure's actual negative impact on price will be higher.

ii. *The Social Welfare Effects of Type I and Type II Error and the "Materiality Threshold."*

— Assessing the social-welfare effects of these tradeoffs between Type I and Type II error starts with understanding both the social benefits and social costs of imposing section 10(b) and Rule 10b-5 liability on an issuer for making a misstatement that inflates an issuer's share price and how the extent of this distortion affects this calculation. Forcing the issuer to pay damages in response to a fraud-on-the-market action helps deter other issuers from making such misstatements in the future. This penalty is the key private mechanism for enforcing the Exchange Act's comprehensive system of mandatory disclosure applicable to publicly traded issuers—if the disclosure is not accurate, damages are assessed.²³⁶ The improved share-price accuracy and issuer transparency that result from greater compliance with this disclosure system increase social welfare by enhancing the efficiency with which resources are allocated in our economy.²³⁷

This gain does not come for free; securities litigation uses scarce resources that could otherwise be deployed to other useful purposes. These resources include the lawyers' and experts' time on both sides of such litigation, as well as the time and effort expended by the

236. See, e.g., John C. Coffee, Jr., *Law and the Market: The Impact of Enforcement*, 156 U. Pa. L. Rev. 229, 266–74 (2007) (detailing methods and results of private enforcement).

237. Greater issuer transparency and more accurate prices in the secondary market signal when an issuer's management is doing a poor job of utilizing the firm's current assets and investing in new projects. Transparency and price accuracy help align the interests of managers with those of shareholders by boosting the effectiveness of share-price-based compensation, block-holder activism, and hostile-takeover mechanisms for reducing the agency costs of management. Greater transparency, by reducing the opportunities for insider trading, also adds to the value of an issuer's shares by increasing their liquidity. These points are discussed in more detail in Fox, *Civil Liability*, *supra* note 124, at 252–60, 264–67 ("By reducing the amount of nonpublic information . . . ongoing periodic disclosure should therefore reduce bid/ask spreads, increase liquidity, and, consequently, reduce the cost of capital.").

issuer's executives and by the judiciary.²³⁸ The amount of resources consumed by such litigation is similar whether the original misstatement resulted in a large or small price distortion. Thus, ideally, liability should be imposed only in cases where, at the margin, the improvement in economic welfare from deterring issuer misstatements is at least as great as the social costs arising from prosecuting the action.²³⁹ This suggests, in turn, that there is some degree of price distortion (as proxied by the size of the corrective disclosure's actual negative effect on price) below which price distorting misstatements should not be subject to fraud-on-the-market damage liability because the costs of the action would exceed the deterrence achieved. We will call this point the "materiality threshold"—the point at which the price distortion is large enough that deterring misstatements that have so significant an effect is worth the enforcement cost.²⁴⁰

iii. *The Meaning of the Assumption that the 95% Confidence Level Is Socially Optimal in Normal Times.* — With this understanding of the social benefits and social costs associated with

238. A reasonable estimate of the litigation costs for the legal and expert fees of both sides is \$2.5 billion annually. See *id.* at 247 n.18 (“[T]his would suggest that the total annual legal expenses associated with these actions averaged about \$2.46 billion . . .”). This figure does not include the value of executive time devoted to defending the litigation or much of the judicial resources consumed by such litigation. Nor does it include the time and resources that honest executives devote to be sure that disclosures that they believe to be true do not, despite the scienter requirement, generate liability because of legal error. See, e.g., Amanda M. Rose, *Fraud on the Market: An Action Without a Cause*, 160 *U. Pa. L. Rev. PENNumbra* 87, 94 (2011), [https://www.law.upenn.edu/journals/lawreview/articles/volume160/issue1/Rose160U.Pa.L.Rev.87\(2011\).pdf](https://www.law.upenn.edu/journals/lawreview/articles/volume160/issue1/Rose160U.Pa.L.Rev.87(2011).pdf) [<http://perma.cc/R5PU-MABT>] (“[I]t is cheap and easy for corporate officers not to lie, but to avoid being *misjudged* a liar, they may spend excessive corporate resources scrubbing disclosures . . .”). The figure also does not include efficiency losses from the decision of managers, out of fear of such liability, not to voluntarily disclose true information before disclosure is otherwise required.

239. This account of the potential social benefits of fraud-on-the-market litigation does not include providing compensation as a way of correcting for unfair investor losses or for an inefficient allocation of risk if investor losses are left where they originally lie. These compensatory rationales for imposing liability simply do not hold up under close examination, a view widely shared by commentators on the issue. See, e.g., John C. Coffee, Jr., *Reforming the Securities Class Action: An Essay on Deterrence and Its Implementation*, 106 *Colum. L. Rev.* 1534, 1556–66 (2006) (arguing most compelling rationale for imposing securities fraud liability is “impact of fraud on investor confidence and thus the cost of equity capital” throughout economy); Paul G. Mahoney, *Precaution Costs and the Law of Fraud in Impersonal Markets*, 78 *Va. L. Rev.* 623, 632 (1992) (“It is therefore not surprising that Easterbrook and Fischel’s ‘net harm’ analysis reaches only tentative conclusions when applied to secondary-market frauds.”). The inadequacy of these compensatory rationales is explored in detail in Fox, *Securities Class Actions*, *supra* note 133, at 1192–99 (concluding “investor protection arguments for imposing liability on an issuer to deter misstatements are weak”).

240. What level the materiality threshold should be set at is subject to differences in opinion. These differences arise from different assessments of a variety of factors: the power of fraud-on-the-market suits to deter, the importance of finely accurate share prices, and the total social costs of such litigation. The same differences in these assessments lead to different views on the value of fraud-on-the-market suits as recently rehearsed in the various opinions in the Supreme Court’s decision in *Halliburton Co. v. Erica P. John Fund, Inc.*, 134 S. Ct. 2398 (2014).

imposing liability for misstatements made with scienter and how they relate to the misstatement's *actual* effect on price, we can now explore the meaning of the assumption that, at least in normal times, the 95% standard constitutes the socially optimal set of tradeoff points between Type I and Type II errors.

Type I error—with its resulting imposition of liability where the actual price impact of the corrective disclosure is zero—is the easy case: It unambiguously reduces social welfare. Litigation is costly and there is no gain in deterring misstatements so unimportant that their corrections have no effect on price.

Assessing the welfare impact of Type II error is slightly more complicated. Consider first a misstatement whose actual impact on price is greater than the materiality threshold. Type II error with respect to whether its corrective disclosure had an actual negative effect on price *reduces* social welfare because the error results in a failure to impose liability in a situation where, by definition, imposing liability would have been socially desirable: The improvement in economic welfare from the issuer misstatements that would have been deterred would have been greater than the social costs of the legal action necessary to impose liability. For a misstatement whose actual impact on price is less than the materiality threshold, however, the opposite is the case. Type II error, by blocking imposition of liability in a situation where the social benefits are less than the social costs actually *increases* social welfare.

The social impact of the chosen level of statistical confidence, therefore, needs to account for three factors: the social harm from its Type I error, the social harm from its Type II error with respect to the corrective disclosures of misstatements having actual price impacts greater than the materiality threshold, and the social gain (or social costs avoided) from its Type II error with respect to the corrective disclosures of misstatements having actual price impacts smaller than

the materiality threshold. Given these three factors, consider what is implied by the assumption that the 95% confidence level is socially optimal—that making the confidence standard stricter would reduce social welfare—in normal times.²⁴¹ That is, upon requiring a stricter confidence level, the social-welfare loss from the increase in Type II error with respect to misstatements having price impacts greater than the materiality threshold (factor 2) would exceed the combined social-welfare gain from the decrease in Type I error (factor 1) and the increase in Type II error with respect to misstatements having price impacts less than the materiality threshold (factor 3). And it means that the net social-welfare impact of making the standard less strict would also be negative: The total losses from the increase in Type I error (factor 1) and from the decrease in Type II error with respect to misstatements having price impacts less than the materiality threshold (factor 3) would exceed the gain from the decrease in Type II error with respect to misstatements having price impacts greater than the materiality threshold (factor 2). If 95% is the socially optimal confidence level, requiring either a stricter or a laxer confidence level reduces social welfare.

3. *The Implications of Increased Idiosyncratic Risk for the Use of Event Studies in Fraud-on-the-Market Suits.* — As noted above, an increase in idiosyncratic risk leads to a worsening of the terms of tradeoffs between the Type I and Type II errors. Recall the example where the standard error increases from 1.78%, which was the average standard deviation for the idiosyncratic volatility of the typical firm from the 1970s up until the financial crisis, to 3.23%, the average such standard deviation at the height of the financial crisis.²⁴² Recall that the cutoff—the minimum drop in the market-adjusted price that meets the 95% confidence

241. That is, assuming social welfare will be lower if the chosen confidence level were either stricter or less strict than 95%.

242. See *supra* section III.B.1.b (discussing example); *supra* note 205 and accompanying text (discussing standard deviation at height of 2008 to 2009 financial crisis).

standard—jumps from -3.5% in normal times to -6.33% in crisis times. Maintaining the 95% standard during a crisis-driven high-volatility period means that for corrective disclosures with an actual price impact of -5% , the Type I error rate remains at 2.5%, but the Type II error rate more than triples from about 20% to 66%. The consequences of this worsening tradeoff for fraud-on-the-market litigation are dramatic: Liability would be imposed in only about one case in three for misstatements whose corrective disclosures have a -5% actual impact on price, compared with four cases out of five in normal times.²⁴³

More generally, with idiosyncratic volatility for a firm's stock at a normal level, the choice of the confidence level is the choice of a set of points on the tradeoff between Type I and Type II error, each point corresponding to a particular magnitude of a corrective disclosure's actual negative impact on price. If idiosyncratic risk increases, the terms of this tradeoff worsen. So, for a given chosen level of confidence and hence given Type I error rate, the Type II error will be higher for any given actual negative impact on price.

In the discussion that follows, we consider whether this sharp worsening of the Type I/Type II error tradeoff in high-idiosyncratic-volatility times suggests needed changes of law or policy.

The discussion starts by exploring whether the apparent problem is in fact self-correcting. Some kinds of corrective disclosures will cause larger price drops during crises than normal

243. Here and elsewhere in our discussions of the effect of idiosyncratic risk on event studies, we use a disclosure event with an actual impact on price of 5% as our example for comparative calculations of Type II error rates in normal and in crisis times. Five percent was chosen because a common rule of thumb used by securities lawyers is that information relating to a change in net income of 5% or more is considered material. See *United States v. Nacchio*, 519 F.3d 1140, 1162 (10th Cir. 2008) (“[A] 5% numerical threshold is a sensible starting place for assessing the materiality of [the alleged misstatements], but it does not end the inquiry. Special factors might make a smaller miss [of reported financial performance] material.”), vacated in part on other grounds on reh’g en banc, 555 F.3d 1234 (10th Cir. 2009); see also SEC Staff Accounting Bulletin No. 99, 64 Fed. Reg. 45150, 45151 (Aug. 19, 1999) (stating SEC staff has no objection to registrants using this “rule of thumb” in this fashion). A 5% decrease in the expected value of all future cash flows, discounted to present value at the discount implied in the market price before the decrease in expectations, would, in an efficient market, result in a 5% decrease in price. The policy-derived materiality threshold discussed here actually may well be above or below 5% of net income but in the abstract, most people would agree that a misstatement that causes a 5% change in stock price warrants attention.

times, thus keeping pace as the cutoff for statistical significance grows during crises. For these types of disclosures, unambiguously no adjustment is needed.

The market's reaction to other kinds of important corrective disclosures, however, will not grow in crises. Thus, we go on to address directly the question of whether the worsened terms of the tradeoff between Type I and Type II errors warrants reducing the required level of statistical confidence to something below 95%. In other words, would it not make sense to accept a few more Type I errors to (partially) counteract the rapid increase in Type II errors during crises?

While this idea has intuitive appeal, the discussion below shows that the welfare effects are in fact ambiguous. In one knife-edge case, reducing the confidence level will improve welfare. In all other cases, reducing the standard during crises will be helpful for some kinds of disclosures, but harmful for others, as compared to reducing the standard during normal times. Because it is unclear whether moving down from the 95% level will in fact be helpful in crises, we recommend a different kind of solution. In high-idiosyncratic-volatility times, the 95% standard should be maintained but, to compensate for the diminished effectiveness of private fraud-on-the-market type enforcement, reliance on, and resources for, SEC enforcement actions should be increased.

a. Is the Problem with Increased Idiosyncratic Volatility Self-Correcting? — Is it possible that the very event—economic crisis—that leads to the increase in idiosyncratic volatility also renders it inconsequential in terms of the worsening tradeoff between Type I and Type II error? The idea is that the crisis raises not only the size of the negative price change necessary to be statistically significant but also the size of the price drop from any corrective disclosure. Such a magnified impact is consistent with our explanations of crisis-induced increased idiosyncratic

volatility: current news becoming more important relative to older news,²⁴⁴ news about the quality of management becoming more important,²⁴⁵ and increased ignorance concerning what facts are relevant.²⁴⁶ To see this, suppose that any one or more of these explanations is correct. Then corrective disclosures with actual negative impacts on price that, in normal times, would usually be accompanied by observed market-adjusted price changes sufficiently negative to meet the normal-times cutoff for the 95% normal-times confidence level (-3.49% in our example²⁴⁷) will, in crisis times, have an actual impact on price that is sufficiently more negative that their accompanying observed market-adjusted price changes would be sufficiently negative to meet the 95% confidence level's more restrictive economic-crisis-times cutoff (-6.33% in our example²⁴⁸). For such misstatements, there is no need to worry about a policy response to the worsening of the terms of tradeoff between Type I and Type II errors because the corresponding magnification of the actual negative price impacts of their corrective disclosures erases the effect.

While self-correction of this type will indeed help with regard to certain misstatements, a problem remains. Recall that the most convincing social benefit from allowing fraud-on-the-market causes of action is that they deter price-distorting misstatements and so enhance allocative efficiency in the real economy.²⁴⁹ The absolute level of the distortion and how it compares to what we call the materiality threshold is what is important here, rather than how the distortion compares with some elevated level of overall idiosyncratic price volatility.

244. See *supra* section II.A.

245. See *supra* section II.B.

246. See *supra* section II.C.

247. See *supra* section III.B.1.b (discussing Type II errors).

248. See *supra* section III.B.2.d (discussing effect of increase in idiosyncratic volatility).

249. See *supra* notes 236–238 and accompanying text.

There are two types of misstatements that in crisis times would move prices enough to meet the materiality threshold, but that are not of a nature where the self-correction described here would offset the crisis-elevated cutoff for 95% statistical significance. The first is where the misstatement's actual impact on price in normal times meets the materiality threshold²⁵⁰ but this impact is not magnified by crisis. A statement by an issuer that it holds assets, such as oil reserves, that in fact do not exist is an example. The information-based explanations that we find most plausible for crisis-induced spikes in idiosyncratic volatility²⁵¹ do not suggest that the actual price impact of such a misstatement would be magnified during crisis times, and so it would distort price no more in crisis than in normal times.²⁵² As a result, in crisis times, the price effect of corrective disclosure would be significantly less likely to meet the 95% confidence level's more stringent crisis-period cutoff and so the Type II error rate for this kind of statement will increase. The social benefit from deterring these kinds of misstatements, however, will still be as great as it is in normal times.

The second kind of misstatement where self-correction would not suffice is the flip side of the first: a misstatement that, if made in normal times, would not distort prices sufficiently to reach the materiality threshold but if made in crisis times, is magnified in its price impact sufficiently to reach the materiality threshold. With this magnified impact on price, it becomes socially beneficial to impose liability when it would not be socially beneficial in normal times. In

250. See *supra* note 240 and accompanying text (defining “materiality threshold” as degree of share-price distortion large enough that value of deterring misstatements responsible for distortion exceeds social cost of fraud-on-the-market litigation required to deter it).

251. See *supra* Part II (postulating explanations that current news becomes more relevant during economic crises, information concerning quality of management becomes more relevant during crises, or crises cause uncertainty as to what information is even relevant to begin with).

252. Many disclosures will contain aspects that are “self-correcting” and others that are not. For example, if a company misses its earnings target by \$1 per share, this can be thought of as two pieces of news: (a) shareholders have \$1 less than they expected in assets and (b) earnings may be on a lower trajectory in the future. We expect the reaction to (b) to be amplified during the crisis but that part (a) will be evaluated the same way in both regular and crisis periods.

crisis times, however, the higher Type II error would reduce the likelihood that such misstatements would trigger liability.

b. Does Increased Idiosyncratic Risk in Crisis Times Call for Reducing the Standard of Statistical Significance? — As we have just seen, for many important kinds of misstatements, crises will radically increase the number of Type II errors if we hold the number of Type I errors fixed. For the average firm, more than two-thirds of corrective disclosures causing a 5% drop in the stock price would be missed if we keep the 95% level during the most recent crisis. Even if, as we assume, the 95% confidence level is the socially optimal standard of statistical significance in normal times, maintaining the 95% standard during crisis times will result in a large increase in Type II errors. In essence, the question is whether any social gains from moderating the increase in Type II errors would outweigh the social losses from the increase in Type I errors. As we will see below, it is impossible to even make an educated guess as to the answer. It is just as likely that reducing the standard in crisis times would decrease social welfare as increase it.

The analysis makes three principal points:

1. To determine whether it would be desirable to relax the standard below the 95% confidence level in crisis times, the focus must be on whether, at the margin, the welfare gains from reduced Type II errors are greater when the standard is more relaxed in crisis times than in ordinary times. This is because the corresponding increase in Type I error will be the same in crisis times as in normal times and so the resulting social losses will be the same.

2. Reducing the standard in crisis times will decrease Type II error by more than doing so in normal times for corrective disclosures with actual negative impacts on price greater than a particular magnitude (the “crossover point”), but will decrease Type II error by less than doing so in normal times where the actual negative price impact is less negative than this crossover point.

3. Except in the unlikely event that this crossover point coincidentally equals the materiality threshold,²⁵³ it will be very difficult to tell whether the welfare gains from reduced Type II errors are greater when the standard is relaxed in crisis times, compared to relaxing it in ordinary times. This is because there will be a range of negative price impacts where it is socially desirable to impose liability and where the reduction in Type II error is greater when the standard is relaxed in crisis times than in normal times. But there will also be a range of negative price impacts either where it is socially desirable to impose liability and the reduction in Type II error is instead less when the standard is relaxed in crisis times than in normal times, or where it is socially *undesirable* to impose liability and the reduction in Type II error is greater when the standard is relaxed in crisis times than in normal times.

i. *A Critical Question.* — The assumption that the 95% confidence level is socially optimal in normal times means that, at the margin, the social loss from a small relaxation in the standard, with the resulting tiny increase in Type I error, *just equals* the social gain from the corresponding decrease in Type II error. It also means that the social loss from a more than infinitesimal increase in Type I error must be greater than the increase in social benefits from the corresponding decrease in Type II error. Hence, the assumption implies that in normal times there would be no improvement from a meaningful relaxation of the confidence standard. If an improvement were possible, the 95% standard would not be optimal.

This analysis means that lowering the required confidence level below 95% can enhance welfare in crisis-induced, high-idiosyncratic-volatility times only if—contrary to normal times—the social gain from a decrease in Type II error *is greater* than the social cost from the corresponding increase in Type I error. Because the confidence level is defined in terms of the

253. That is, the minimum size of price drops from corrective disclosures that it is beneficial to deter with liability.

acceptable level of Type I errors, lowering the standard by a given amount in normal times and in high-idiosyncratic-volatility times results in the same increase in Type I error rate. So lowering the confidence level will have the same impact on Type I errors and the same negative effect on social welfare in both crisis and normal times.²⁵⁴ Thus, the critical question is whether, at the margin, reducing the standard in crisis times will result in greater social welfare gains from reduced Type II errors than reducing the standard in normal times would.

In sum:

1. Given the assumption that the 95% confidence level is socially optimal, if the level is relaxed in normal times, the social gain from the reduced Type II errors at the margin just equals the social cost from the increased Type I errors.

2. The social cost from the increased Type I errors if the standard is relaxed in crisis times equals the social cost from the increased Type I errors from relaxing the standard in normal times.

3. Thus, if the standard is relaxed in crisis times and the social gain from the reduction in Type II errors is greater than if it is relaxed in normal times, then the social gain from the reduction in Type II errors is greater than the social cost from the increase in Type I errors. This would mean that relaxing the standard in crisis times is a good thing to do.

ii. *The Type II Error Effects from Reducing the Confidence Level in Crisis Times Versus Normal Times and the Crossover Point.* — For corrective disclosures with an actual negative impact on price greater than a particular magnitude (the crossover point), reducing the required confidence level from 95% to some lower level in crisis times will decrease Type II error by

254. We are assuming, plausibly, that the costs and benefits of imposing liability for misstatements with any given price impact are the same in crisis and normal times.

more than doing so in normal times would. Where the actual negative price impact is less negative than this crossover point, however, doing so in crisis times will decrease Type II error by *less* than doing so in normal times would. The discussion that follows summarizes the reasoning as to why.²⁵⁵

The first part of this proposition—relating to price impacts more negative than the crossover point—would seem to make intuitive sense. A crisis-induced increase in idiosyncratic risk increases Type II error for any level of price drop, so one would think that reducing the standard would be more effective at reducing Type II error when there was more Type II error to reduce. For the same reason, the second part of the proposition—relating to price impacts less negative than the crossover point—at first seems counterintuitive.

An extreme example suggests why both parts of the proposition in fact make sense. Imagine that in crisis times, the standard deviation of the typical firm's idiosyncratic volatility is 10%, whereas in normal times, it is 1%. Now compare the reduction in Type II error from lowering the confidence level in normal times versus crisis times where the actual negative price impact of the corrective disclosure is a large 20%. During normal times Type II errors for this kind of disclosure are essentially zero. This is because the net impact of other firm-specific news and background noise on the day of the disclosure is almost never so positive as to disguise the actual 20% drop: The assumed 1% normal-times standard deviation of idiosyncratic risk tells us that the net impact of these other items will rarely move the price more than a few percent. So the Type II error rate with the 95% standard is almost zero. If the standard is lowered to 90%, the Type II error rate will be even closer to zero, but the reduction in Type II error from reducing the standard will be very small.

255. We have presented a rigorous proof of this proposition in a different paper. Fox, Fox & Gilson, *Idiosyncratic Risk*, *supra* note 186, at 11–19.

In contrast, in crisis times, the standard deviation of 10% tells us that the net impact of other firm-specific news and background noise will move price much more relative to this 20% drop. The Type II error rate with the 95% standard would be 48%, whereas if the standard were relaxed to 90%, it would be 36%. Thus, for a corrective disclosure with a large actual price effect, reducing the confidence level results in a much larger reduction in Type II error in crisis times than it would in normal times.

Where the negative price impact of the corrective disclosure is a much smaller 2%, the opposite will be the case. In normal times, the standard deviation of 1% tells us that other firm-specific news and background noise will move price much more relative to an actual 2% negative price impact than to an actual 20% price impact. In this normal period, the Type II error rate with the 95% standard would be 48% and with the 90% standard it would be 36%. In contrast, in crisis times, an actual negative price impact of 2% is a drop in the bucket compared to the swings expected from the net impact of other firm-specific news and background noise. Type II error thus will be very large whichever confidence level is used. The results of an event study would be statistically significant at the 95% level only if the observed market-adjusted price change was -19.6% or more, and statistically significant at the 90% level only if it was -16.4% or more. To pass these tests, the net impact of other firm-specific news and background noise would have to be at least -17.6% and -14.4% , respectively. The net impact of these other items is slightly more likely to reach -14.4% or -17.6% or more during a crisis, but the difference in likelihood is small. Specifically, the Type II error with the 95% confidence level standard would be 96%. With a 90% confidence level, the Type II error rate would be 92%. Thus, for a corrective disclosure with a small actual price effect, reducing the required

confidence level results in a much smaller reduction in Type II error in crisis times than it would in normal times.

Keeping in mind this extreme example, consider a more realistic example. Assume a corrective disclosure with an actual negative price impact of 5%. Recall that the observed market-adjusted price change will be the combination of this 5% drop and the net effect of the other bits of firm-specific news and background noise that affect the issuer's share price the same day.²⁵⁶ The net impact of these other items is, on average, zero. So the distribution of possible observed market-adjusted price changes will be centered around -5%. How widely the returns are dispersed around -5%, however, is a function of the firm's idiosyncratic volatility: the higher the volatility, the wider the dispersion. Figure 11(a) depicts two dispersions around -5% of possible observed market-adjusted prices for the corrective disclosure in our example. The shallower curve represents a standard deviation of 3.23%, the standard deviation of the typical firm in the recent financial crisis. The steeper curve represents a standard deviation of 1.78%, the standard deviation of the typical firm in normal times.

256. See *supra* section III.B & fig. 6 (studying statistical causal link between factors such as public firm-specific information and observed price change).

**Figure 11(a) When the Corrective Disclosure Causes A 5% Drop in Price
—Comparing Regular and Volatile Periods**

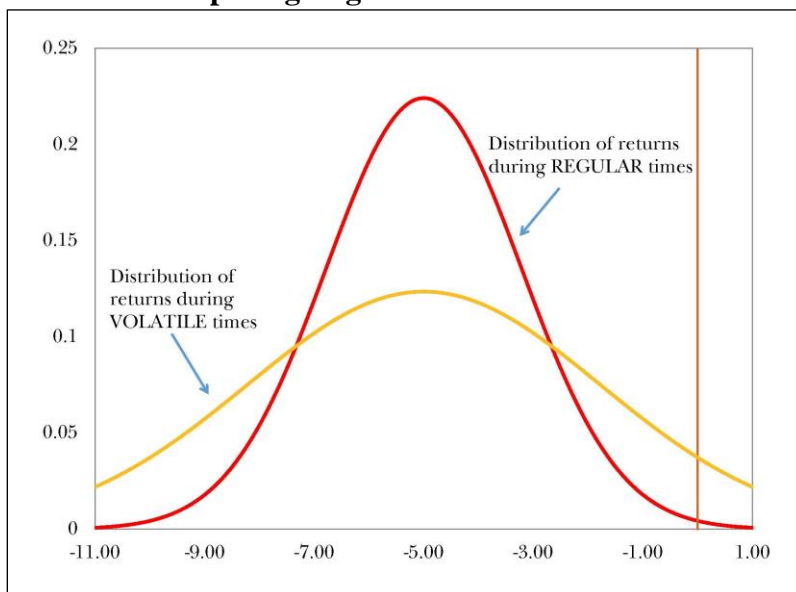


Figure 11(b)
5% DROP—TYPE II ERROR USING 95% LEVEL IN
VOLATILE PERIODS

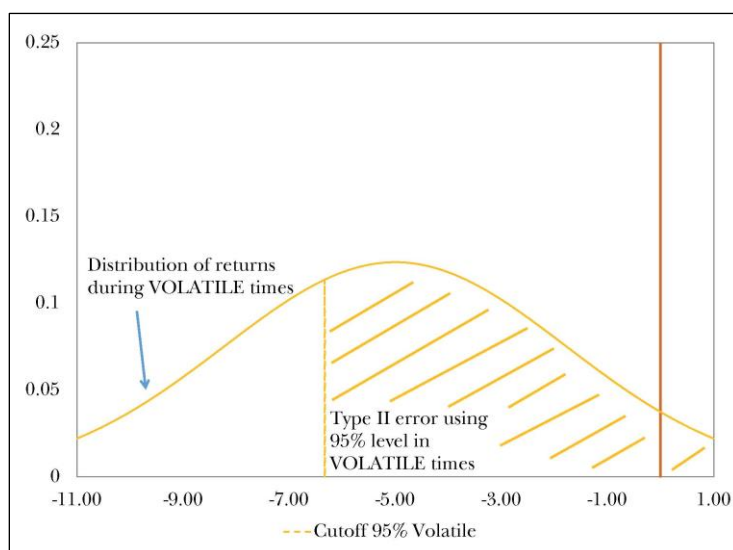


Figure 11(c)
5% DROP—TYPE II ERROR USING 95% LEVEL IN
NORMAL PERIODS

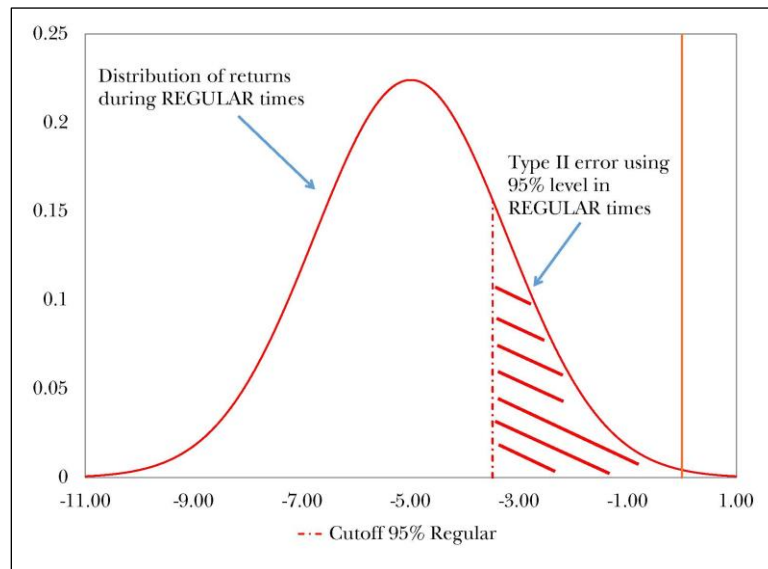


Figure 11(d)
5% Drop—Reduction in Type II Error Moving To 90% Level in Volatile Periods

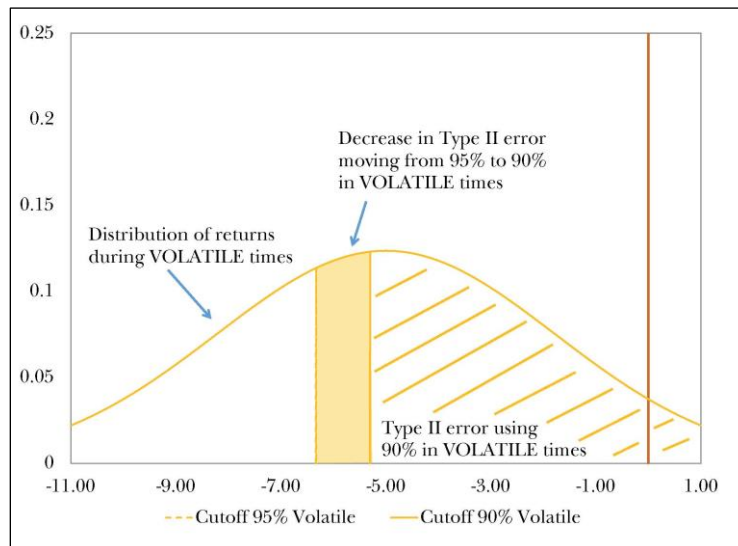


Figure 11(e):
5% Drop—Reduction in Type II Error Moving To 90% Level in Regular Times

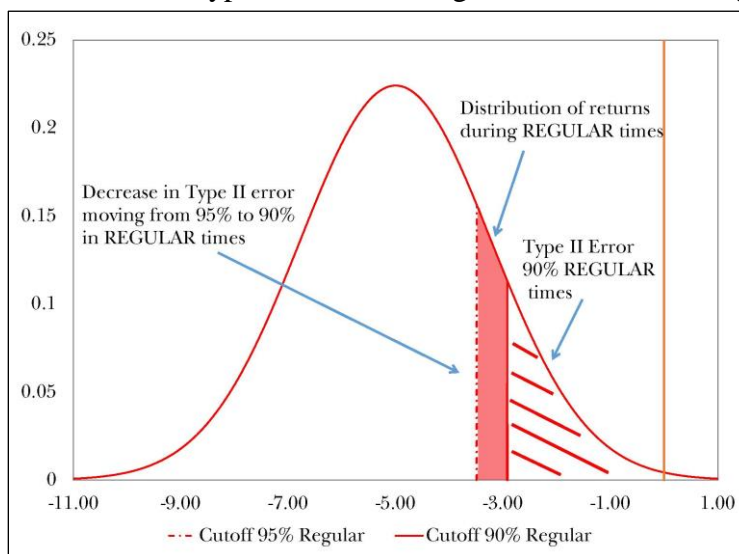
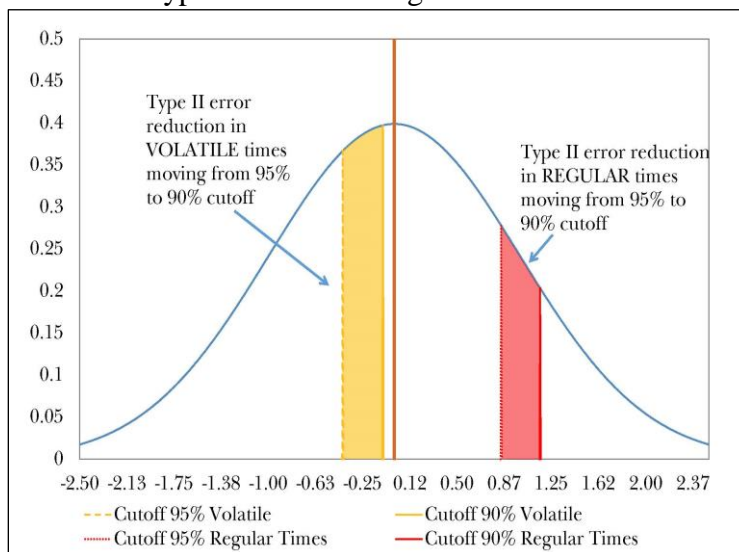


Figure 11(f)
5% Drop—Comparing Reductions in
Type II Error Moving To 90% Level



Now recall that the firm's idiosyncratic volatility also determines how negative the observed market-adjusted price change must be for it to be considered statistically significant at the 95% confidence level.²⁵⁷ The cutoff is -1.96 multiplied by the standard deviation of the idiosyncratic

257. See *supra* section III.B.1.b (mentioning relevance of firm-specific volatility on observed price change).

risk of the firm. So for this typical firm, this cutoff in the financial crisis was $-1.96 \times 3.23 = -6.33\%$ and in normal times would be $-1.96 \times 1.78 = -3.49\%$. Type II error—a corrective disclosure with an actual negative impact on price that fails the test—occurs when the observed market-adjusted price change is not sufficiently negative to meet the cutoff.

Figures 11(b) and 11(c) display the Type II error rate using the distributions from 11(a) and the crisis- and normal-period cutoffs, respectively: In the figures, any observed market-adjusted price to the right of the applicable cutoff fails the test. For example, in Figure 11(b), because -5% is less negative than the 95% cutoff for crisis periods (-6.33%), the corrective disclosure in our example will only pass the test if the net impact of other bits of firm-specific news and background noise drives the price down by at least another 1.33% . This is 0.41 standard deviations below the distribution mean (which is -5%). For a normal distribution, this will happen only about one-third of the time, so the observed price will only be negative enough to pass the test about one-third of the time. As a result, Type II error in crisis periods will be about 66% (i.e., the area marked under the curve to the right of the cutoff in Figure 11(b) is 66%).

In contrast, as depicted in Figure 11(c), during normal times, the observed market-adjusted price will pass the test unless the net impact of these other items is sufficiently positive to drive the observed price up by at least 1.51% . This is 0.84 standard deviations above the mean, which with a normal distribution occurs only about 20% of the time (i.e., the marked area in Figure 11(c) under the curve to the right of the cutoff point is 20%). In normal times, then, Type II error would only be 20% .

Figures 11(d) and 11(e) graphically show what happens when we relax the statistical standard to 90% . The cutoffs then drop to -5.30% ($-1.64 \times 3.23\%$) during crisis times and 2.92% ($-1.64 \times 1.78\%$) in normal times. The area under the curves in these two figures between

the 95% and 90% thresholds represents the reduction in Type II errors. The area in Figure 11(d) is larger than in Figure 11(e), indicating that a corrective disclosure with an actual negative impact on price of -5% is above the cutoff point: Type II errors decline after reducing the required confidence level by more during crisis times than during regular times.

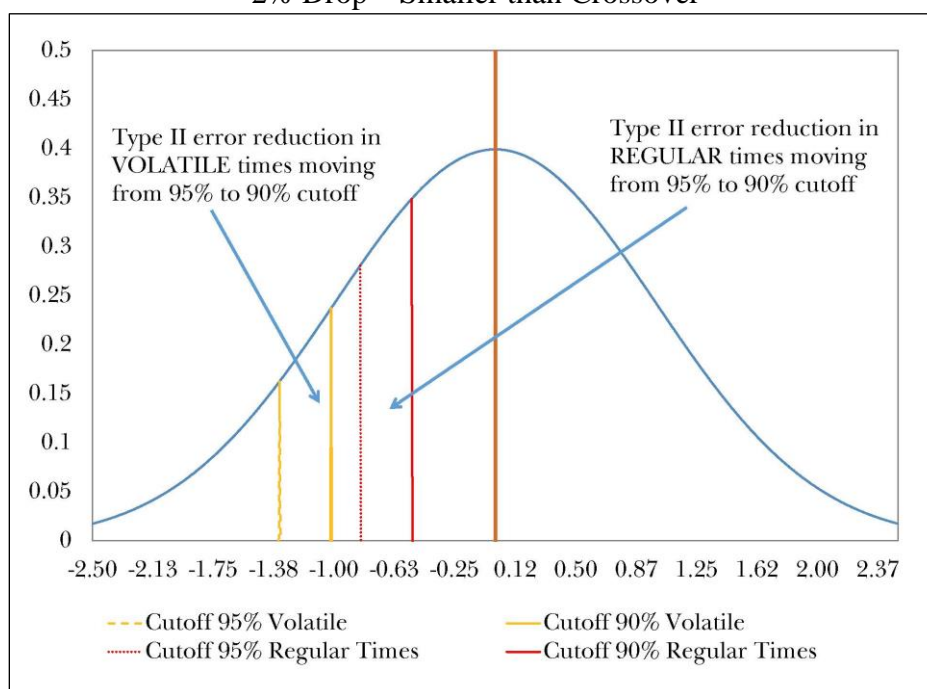
Figure 11(f) displays the same information as Figures 11(d) and 11(e), except that we have transformed the two distributions so they can be directly compared by turning them both into the standard normal distribution, centered at zero and having a standard deviation of 1.²⁵⁸ Observe two things. First, this transformation preserves the standard deviations we noted above, so that the threshold for 95% in crisis times starts 0.41 standard deviations below the mean and the 95% threshold in normal times starts 0.84 standard deviations above the mean. Second, observe that the interval between the 95% and 90% cutoff is the same length (.32) in both regular and crisis periods. This means comparing the reduction in Type II errors is easy: Whichever period's interval is closer to zero (with a -5% actual impact, the crisis period) will see the larger reduction in Type II errors, since the normal distribution is at its highest at zero.

This approach, depicted in Figure 11(f), can be used to compare the normal- versus crisis-period reduction in Type II error from reducing the confidence level for corrective disclosures with any magnitude of actual negative impact on price. Figures 12(a), (b), (c), (d), and (e) depict these comparative reductions for actual negative impacts of 2%, 3%, 4.13%, 5%, and 6%, respectively. One can see from these figures that it will also always be the case that the crisis interval falls to the left of the normal-times interval (because the actual price impact relative to the standard deviation is always less negative in crisis times). One can see from Figure 12(c) that the crossover point will be where the two intervals are mirror images of each other, with the

258. The math, though a bit difficult, is not important here.

crisis-times interval on the negative side and the normal-times interval on the positive side. In all these curves, the distance under the curve from -1.96 to -1.64 (i.e., -0.32) is the same. For actual impacts (Figures 12(d) and 12(e)) that are more negative than the crossover point of -4.13% ,²⁵⁹ the curve is higher over the interval representing the reduction in a crisis period, representing a larger area and hence a larger reduction in Type II error. For actual impacts below the crossover point (Figures 12(a) and 12(b)), the opposite is the case.

Figure 12(a): Visualizing the Reduction in Type II Error in Altering Threshold
2% Drop—Smaller than Crossover



²⁵⁹ The crossover point for relaxing the standard from the 95% confidence level to the 90% confidence level (i.e., allowing up to 5% Type I error instead of only 2.5% Type I error) when volatility increases for the typical company from a standard deviation of 1.78 to 3.23 (the normal times versus the financial crisis) is -4.13% . For the calculations, see Fox, Fox & Gilson, *Idiosyncratic Risk*, *supra* note 186, at 16–19.

Figure 12(b):
3% Drop—Smaller than Crossover

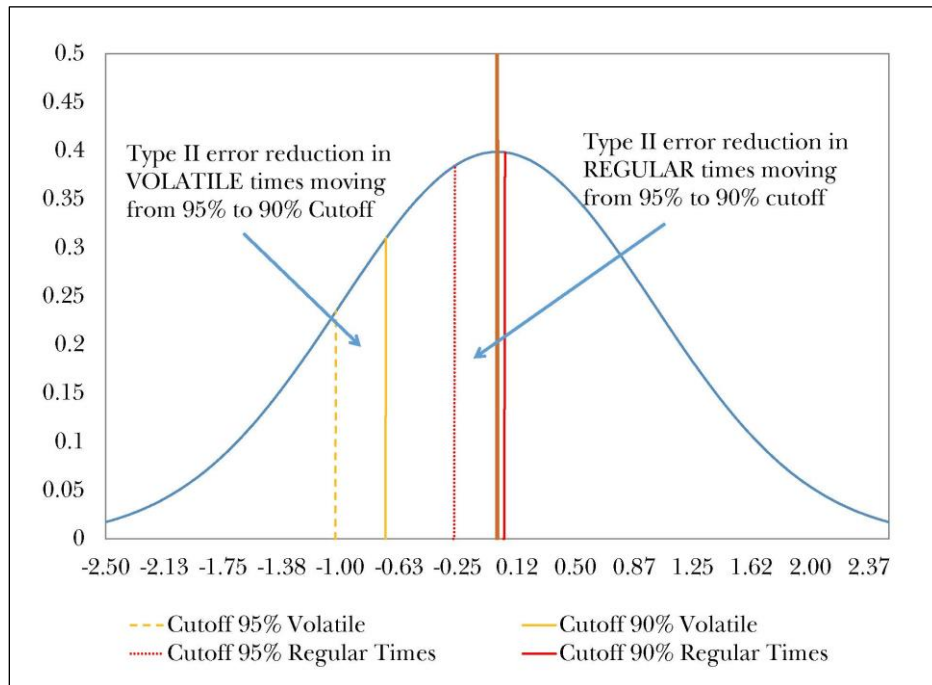


Figure 12(c)
4.13% Drop—Crossover Point

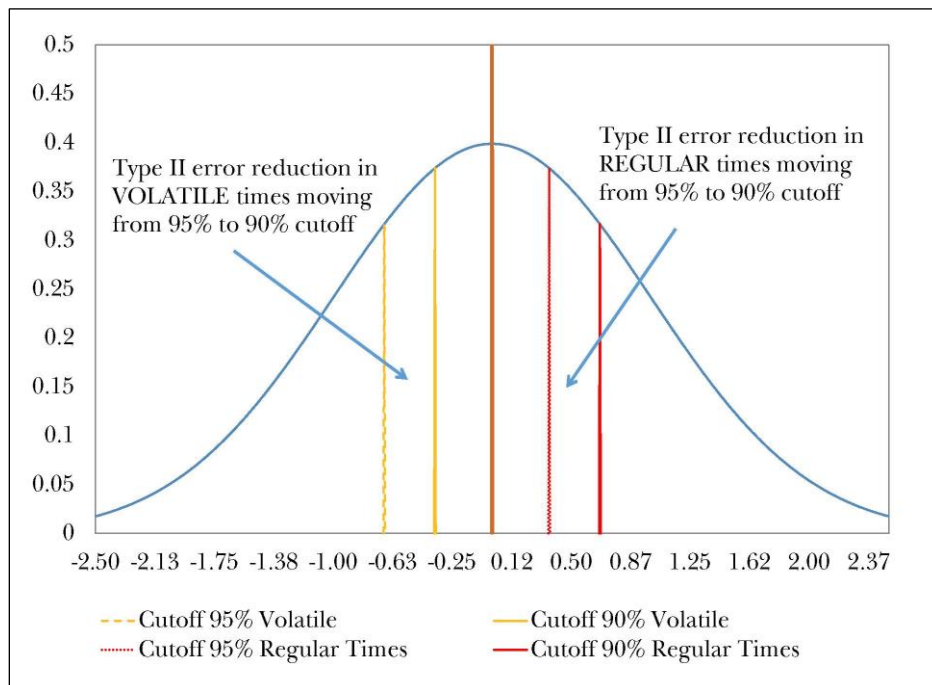


Figure 12(d):
5% Drop—Larger than Crossover

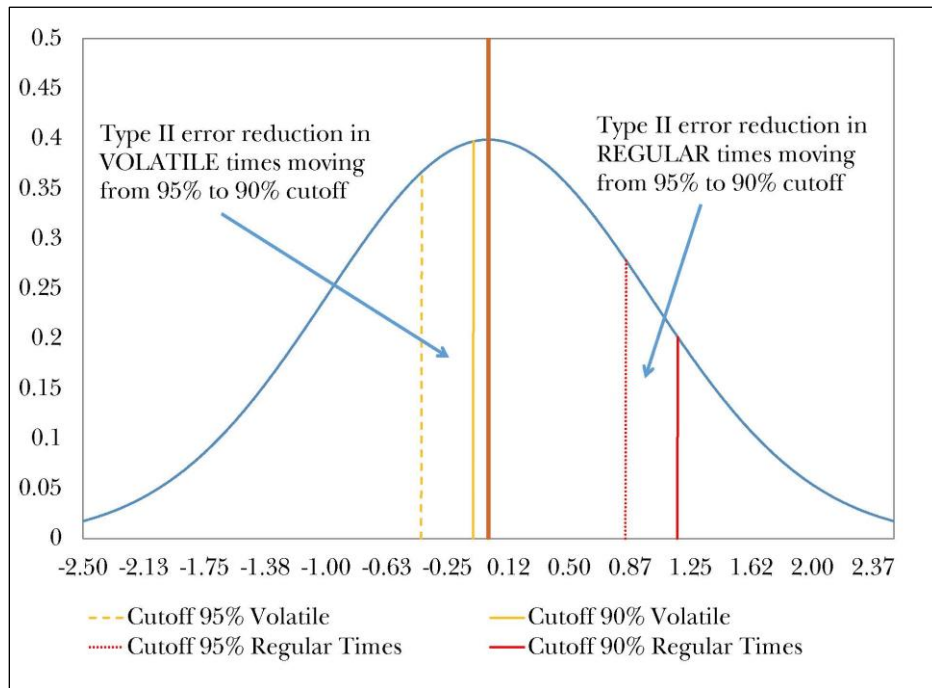
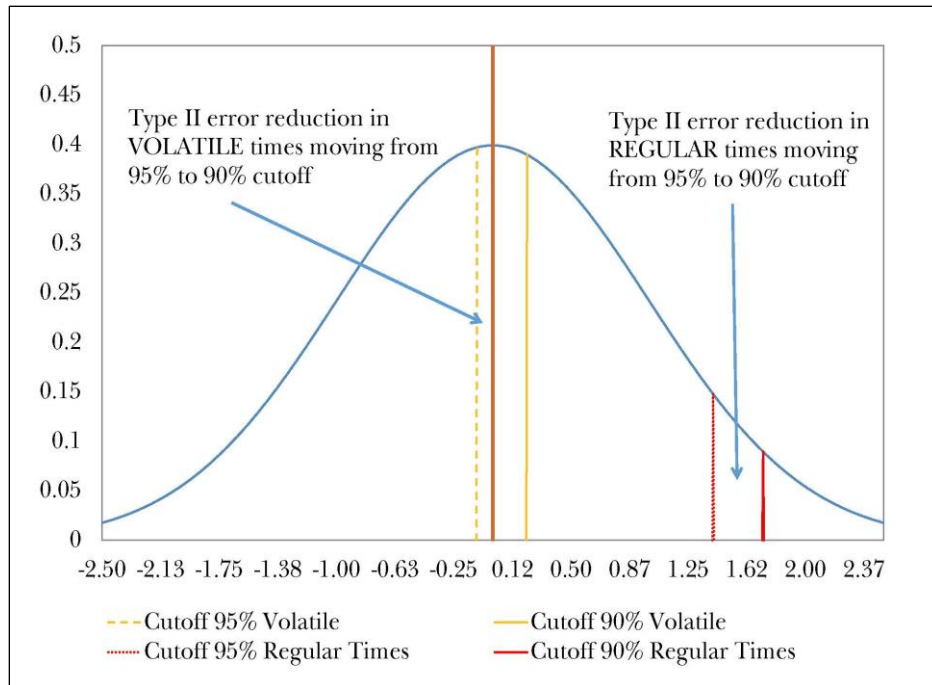


Figure 12(e):
6% Drop—Larger than Crossover



iii. *The Special Case Where the Crossover Point Exactly Equals the Materiality Threshold.*

— To assess the welfare impact of lowering the confidence level in crisis times, consider first the special case where the Type II error crossover point exactly equals the materiality threshold, as depicted by the first bar in Figure 13. In this special case, the impact on Type II errors from reducing the required statistical confidence would have an unambiguously greater net positive effect on social welfare in a period of crisis-induced high idiosyncratic risk than it would in normal times. To see why, recall that it is socially desirable to impose liability on misstatements made with scienter whose corrective disclosures have actual price impacts more negative than the materiality threshold because the social benefits from the deterrent effects of imposing liability exceed the litigation's social costs.²⁶⁰ The converse would be true for misstatements whose corrective disclosures have actual price impacts smaller than the materiality threshold.

For all corrective disclosures with actual price impacts more negative than the crossover point, reducing the confidence level will decrease false negatives more in high-idiosyncratic-volatility times than in normal times. Where the crossover point just equals the materiality threshold, this is exactly the range of price impacts where reducing false negatives is welfare-enhancing. This is because it is desirable for liability to be imposed where the corrective disclosure's price impact is more negative than the materiality threshold.

For all corrective disclosures with actual price impacts less negative than the crossover point, lowering the confidence level in high-idiosyncratic-volatility times will decrease false negatives (Type II errors) by less than in normal times. Reducing false negatives for corrective disclosures with price impacts in this range is welfare-destroying because it is undesirable to impose liability where the corrective disclosure's price impact is less negative than the

260. See *supra* note 240 and accompanying text (explaining concept of “materiality threshold” as point at which price distortion becomes large enough to justify enforcement costs of deterring misstatements).

materiality threshold. So less reduction in false negatives is desirable because false negatives block imposition of liability and imposing liability is not desirable for corrective disclosures with actual negative price impacts this small.

Now we can put the two cases together. As depicted in the first bar in Figure 13, over the full range of possible actual negative price effects from corrective disclosures, the impact on the Type II error rate from reducing the confidence level would increase social welfare in a period of high idiosyncratic risk. Compared to normal times, reducing the confidence level in crisis times would reduce the Type II error rate by more in the range where Type II errors are undesirable and reduce it by less in the range where Type II errors are desirable. Therefore, in this special situation, lowering the standard in volatile times would unambiguously increase social welfare.

iv. *The Ordinary Case Where the Crossover Point Does Not Equal the Materiality Threshold.* — The special case described above, where the crossover point is exactly equal to the materiality threshold, would be pure coincidence because the factors determining each are independent. In the ordinary case, they would not be equal. As a result, the comparative welfare effects of lowering the confidence level in crisis times versus doing so in normal times become more complicated. If the crossover point is either more or less negative than the materiality threshold, there will be a range of actual negative price effects from corrective disclosures for which the impact on Type II errors from lowering the standard in crisis times will have a less positive, or a more negative, effect on social welfare than in normal times.

First consider the situation, depicted in the second bar in Figure 13, where the crossover point is more negative than the materiality threshold. In this situation, for corrective disclosures with actual price impacts less negative than the crossover point but more negative than the materiality threshold, reducing the standard in high-volatility crisis times will reduce false

negatives by less than doing so in normal times. This is a range of actual price impacts where false negatives are undesirable. So for corrective disclosures with price effects in this range, lowering the standard in crisis times is, in terms of its impact on Type II errors, less socially beneficial than doing so in normal times.

Next consider the opposite situation, depicted by the third bar in Figure 13, where the crossover point is less negative than the materiality threshold. For corrective disclosures with actual price impacts more negative than the crossover point but less negative than the materiality threshold, lowering the confidence standard in volatile times will reduce false negatives by more than doing so in normal times. This is a range of actual price impacts, however, where false negatives are desirable because it is undesirable to impose liability. So for corrective disclosures with price effects in this range, lowering the standard is, in terms of its impact on Type II errors, more socially harmful than doing so in ordinary times.

Thus, in each of these two situations, for corrective disclosures with price impacts in the range between the materiality threshold and the crossover point, the welfare effects of reducing the confidence level in high-volatility crisis times would be less beneficial, more harmful, than it would be in normal times. For corrective disclosures that have actual price effects that are on either side of this range, the welfare effects of lowering the standard would be more favorable, or less unfavorable, in crisis times than in normal times, for the same reasons as in the special situation where the crossover point precisely equals the materiality threshold.

If, as seems likely, the crossover point does not equal the materiality threshold—meaning there is a range of corrective disclosure price impacts that fall between these two points—what can we conclude about the social welfare effect of lowering the confidence level during crisis-induced high idiosyncratic volatility times? Recall that the critical question is whether the

welfare benefits from the reduction in Type II errors are greater when the confidence level is relaxed in crisis times compared to ordinary times. Thus, we need to know whether the enhanced level of welfare gains with respect to corrective disclosures with price impacts outside this range on one side or the other dominate the reduced level of welfare gains, or increased level of welfare losses, with respect to corrective disclosures with price effects within this range. Answering this question requires knowing two things. First, one must know the distribution of misstatements in the economy in terms of their price effects (and hence the price effects of their associated corrective disclosures). And second, the answer also requires knowing, for corrective disclosures with each such level of price impact, the social gain or loss arising from weighing the deterrence benefits from imposing liability versus the costs of such litigation. We are currently far from knowing either of these things.

Figure 13: Type II Error Reduction from Reducing the Confidence Level for Statistical Significance in Crisis Times versus Normal Times and the Implications for Social Welfare

For corrective disclosures with actual impacts more negative than the crossover point, reducing the standard in crisis times will decrease Type II error by more than doing so in normal times would. For corrective disclosures with actual impacts less negative than this point, reducing the standard in crisis times will decrease Type II error by less than doing so in normal times would.

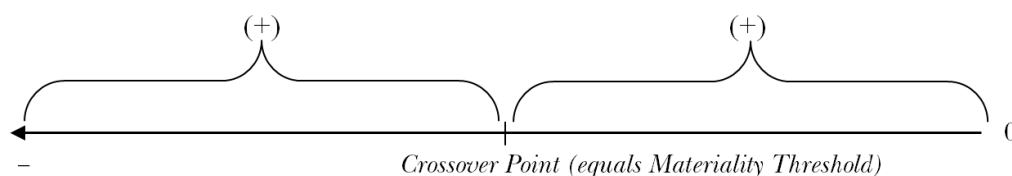
It is socially desirable to impose liability on misstatements made with scienter whose corrective disclosures have actual price impacts more negative than the materiality threshold and socially undesirable to impose liability for such misstatements when they have actual price impacts less negative than this point.

The comparative welfare effects of lowering the standard in normal versus volatile times is depicted here with regard to three situations: (A) where the crossover point equals the materiality threshold, (B) where the crossover point is less than the materiality threshold, and (C) where the crossover point is greater than the materiality threshold.

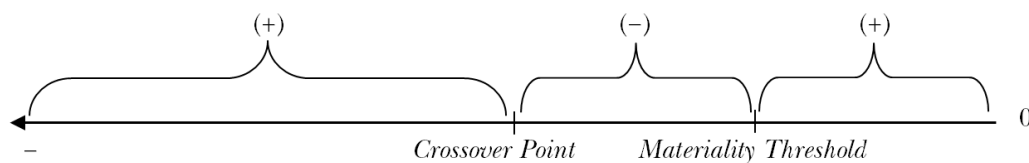
(+) = The welfare gains from reduced Type II errors are greater, or the welfare losses from reduced Type II errors are smaller, in crisis times than in normal times.

(-) = The welfare gains from reduced Type II errors are smaller, or the welfare losses from reduced Type II errors are greater, in crisis times than in normal times.

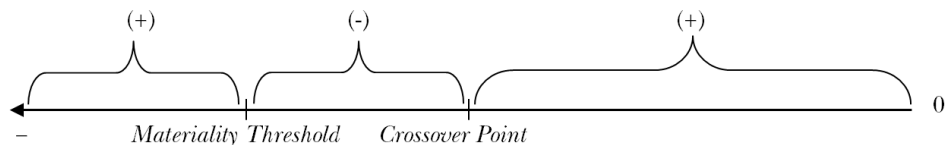
A. Crossover Point equals the Materiality Threshold



B. Crossover Point more negative than the Materiality Threshold



C. Crossover Point less negative than the Materiality Threshold



v. *Summary and Conclusion.* — Whether the steep increase in Type II errors during economic-crisis-induced periods of high idiosyncratic volatility warrants reducing the confidence level used in event studies of corrective disclosures in order to determine loss causation and materiality in fraud-on-the-market suits reduces to the following question: At the margin, does lowering the confidence level in crisis times increase social welfare through the reduction of Type II errors more than doing so in normal times does? This distillation follows logically from two observations. First, the negative welfare effect from a relaxed confidence level's increase in Type I errors will be the same in crisis times as in normal times because, by definition, reducing the standard will increase Type I error by the same amount in each of these two periods. Second, in normal times, this negative welfare effect from a reduced confidence level's increase in Type I errors will, at the margin, just equal the positive welfare effect from the reduced confidence level's decrease in Type II error because this is a necessary condition for our starting assumption that in normal times the 95% standard results in the socially optimal set of points in the tradeoffs between Type I and Type II errors.²⁶¹ Thus, reducing the confidence level in crisis times is desirable only if the resulting positive welfare effect from a reduced standard's impact on Type II errors is greater in volatile times than in normal times.

We have seen that it is impossible to determine, without considerably more information than appears to be available, whether this condition is met. The exception is the special—and purely

261. See *supra* notes 232–234 and accompanying text (explaining assumption that 95% is socially optimal confidence level).

coincidental—case where the Type II error crossover point exactly equals the materiality threshold. That case would be a pure coincidence because the factors determining the crossover point are entirely independent of the factors determining the materiality threshold.

The assumption that the 95% confidence level is the socially optimal point of tradeoff between Type I and Type II errors in normal times may, of course, be incorrect.²⁶² But that would simply suggest that the standard be changed for normal times, an analysis that we do not undertake here (though our discussion identifies the nature of the inquiry). Regardless, the results here are generalizable from any such altered set point. If the optimal standard in normal times is something other than 95%, this analysis still suggests that without more information than we now have, we cannot have strong priors that welfare would be enhanced, rather than compromised, by lowering the required standard of statistical confidence during periods of economic crisis.

c. Shifting the Mix of Enforcement Mechanisms in Crisis Times. — The foregoing discussion shows that there are no simple answers to the questions that arise from using event studies to assess materiality and loss causation in fraud-on-the-market class actions in the face of economic crisis-induced increases in idiosyncratic volatility. We can say with certainty, however, that the dramatic worsening in the tradeoff between Type I and Type II errors makes the threat of fraud-on-the-market actions a comparatively much less useful tool for deterring price-distorting issuer misstatements in crisis times than in normal times. As has just been demonstrated, there is as much reason to believe that lowering the required confidence level in times of economic crisis will decrease social welfare as increase it. And if the required level of statistical confidence is maintained at 95%, the threat of such an action constitutes a considerably less effective deterrent.

262. See *supra* note 233 and accompanying text (discussing assumption).

This loss of deterrence can be illustrated. Assume that if misstatements made with scienter distort prices by 5%, they surpass the materiality threshold, i.e., they are serious enough to incur the social costs of deterring them through private actions. As we have seen from our examples, crisis times would transform this useful cause of action from one that catches most such misstatements—four out of five—to one that catches many fewer—just one in three. And if the confidence standard is lowered to at least partially compensate for this loss in deterrence, a fraud-on-the-market class action becomes a costlier way to deter issuer misstatements because, among other reasons, of the corresponding increase in Type I errors.

This very substantial decline in the usefulness of fraud-on-the-market class actions in crisis times has not been previously recognized. It suggests that during crisis times, more resources should be devoted to other methods for deterring price-distorting misstatements, for example stepped-up SEC enforcement actions. In contrast to fraud-on-the-market suits, the SEC is not required to establish loss causation in Rule 10b-5 enforcement actions.²⁶³ So in a crisis, the SEC will not have to prove loss causation in times when it is considerably less likely that the corrective disclosure of a misstatement that substantially distorts price will be accompanied by a statistically significant, market-adjusted price drop. The SEC still, of course, needs to establish that the misstatement was material, but unlike with respect to loss causation, the case law permits materiality to be proven other than through an event study.²⁶⁴ For example, the SEC could show the facial importance of the issuer's misstatement, the extent to which analysts took note of the misstatement at the time it was made, and any price reaction at the time of the misstatement.

263. See, e.g., *SEC v. Kelly*, 765 F. Supp. 2d 301, 319 (S.D.N.Y. 2011) (“[U]nlike a private plaintiff, the SEC need not allege or prove reliance, causation, or damages in an action under Section 10(b) or Rule 10b-5.”); *supra* notes 209–225 and accompanying text (discussing requirements of fraud-on-the-market shareholder class actions).

264. See, e.g., *SEC v. Lee*, 720 F. Supp. 2d 305, 335 (S.D.N.Y. 2010) (holding SEC sufficiently pleaded materiality by alleging defendant's false statements would have influenced reasonable investor).

Evidence that would suggest that the market-adjusted price drop at the time of any unambiguous issuer announcement of the truth might be smaller than the amount by which the misstatement inflated price also would be relevant. Possible explanations of how this could occur include insider trading based on the truth, rumors of the true situation circulating in the market, and the existence of a series of corporate announcements that dribbled the truth out in small doses in advance of the full corrective disclosure. All of these could lead to the market's realizing the truth, often gradually, in advance of the full corrective disclosure and thus cause a smaller or nonexistent observed market-adjusted price change at the time of the disclosure.²⁶⁵

Using these other indicia to determine whether a misstatement influenced price is a considerably more subjective exercise than the use of an event study. Indeed much of the event study's appeal for the judiciary derives from its greater objectivity and transparency: Experts may differ in their methodology, but in comparison to, for example, the testimony of a financial analyst, the differences between competing methodologies are readily observable. This appeal is so important that in private suits, some courts will grant the defendant summary judgment on the issue of loss causation if the plaintiff does not introduce an event study showing a statistically significant, market-adjusted negative price change at the time of the corrective disclosure.²⁶⁶ But when the event study becomes a comparatively less powerful tool, alternative forms of evidence concerning loss causation and materiality become relatively more attractive. At least in theory, courts could give more importance to these other forms of evidence in fraud-on-the-market suits

265. See Fox, After *Dura*, supra note 222, at 850–51 (explaining why share price does not necessarily drop immediately after corrective disclosure); David Tabak & Frederick C. Dunbar, Materiality and Magnitude: Event Studies in the Courtroom 7 (Nat'l Econ. Research Assocs., Working Paper No. 34, 1999), www.nera.com/content/dam/nera/publications/archive1/3841.pdf [<http://perma.cc/KUK3-FFLF>] (discussing beginning event studies prior to actual announcement of news to correct for leakage before that time).

266. See, e.g., *In re Imperial Credit Indus., Inc. Sec. Litig.*, 252 F. Supp. 2d 1005, 1014–16 (C.D. Cal. 2003) (finding defendants entitled to summary judgment because plaintiffs failed to produce event study); *In re Exec. Telecard, Ltd. Sec. Litig.*, 979 F. Supp. 1021, 1025–27 (S.D.N.Y. 1997) (rejecting reliability of expert testimony due to failure to perform event study).

as well as in SEC enforcement actions in times of economic crisis.²⁶⁷ But it is likely that they are relatively more manageable in the hands of an administrative agency that can use prosecutorial discretion in deciding which cases to bring.

B. Establishing Materiality in Other Securities Law Contexts

The concept of “materiality” pervades securities law, reaching far beyond fraud-on-the-market suits. This includes two particularly important areas where event studies are frequently used: insider trading regulation and section 11 of the Securities Act of 1933 (the Securities Act) litigation based on alleged misstatements or omissions in the registration statements of public securities offerings. Use of event studies to establish materiality in each of these important areas raises sufficiently distinct issues from their use in fraud-on-the-market studies to warrant separate discussion.

1. Insider Trading. — One of the requirements of the longstanding prohibition under Rule 10b-5 against insider trading is that the information on which the insider trades be “material.”²⁶⁸ The doctrinal standard for materiality in insider trading cases under Rule 10b-5 is identical to that in Rule 10b-5 misstatement cases: whether there is a substantial likelihood that a reasonable investor would consider the nonpublic information on which the insider traded to be important in a decision whether to purchase or sell the security involved.²⁶⁹ If the later public disclosure of this information changes the price of this security, it has had an actual effect on the behavior of

267. For a view of one of the authors that as a general matter, some of these factors should, under certain circumstances, be taken into account by courts in fraud-on-the-market actions, see Merritt B. Fox, *Demystifying Causation in Fraud-on-the-Market Actions*, 60 Bus. Law. 507, 523–25 (2005). For an opposing view that they never should be, see John C. Coffee, Jr., *Causation by Presumption? Why the Supreme Court Should Reject Phantom Losses and Reverse Broudo*, 60 Bus. Law. 533, 537 (2005).

268. See, e.g., *SEC v. Tex. Gulf Sulphur Co.*, 401 F.2d 833, 848 (2d Cir. 1968) (en banc) (“An insider’s duty . . . to abstain from dealing in his company’s securities arises only in ‘those situations . . . which are reasonably certain to have a substantial effect on the market price of the security if . . . disclosed.’” (quoting Arthur Fleischer, Jr., *Securities Trading and Corporate Information Practices: The Implications of the Texas Gulf Sulphur Proceeding*, 51 Va. L. Rev. 1271, 1289 (1965))).

269. See *supra* note 224 (explaining definition of materiality).

investors. Again, this change in price strongly suggests that a reasonable investor would have found the information important at the time that the defendant traded on it. As in fraud-on-the-market cases, an event study is the financial economist's standard tool for determining the likelihood that the disclosure of a previously nonpublic item of news in fact did change the price of a security. Not surprisingly, therefore, the courts in insider trading litigation have accepted event studies as at least one (preferred) way to establish the materiality of information.²⁷⁰

The social gains from deterring trades based on inside information are somewhat different from the gains from deterring corporate misstatements, as are the costs from imposing liability for insider trading where the importance of the information is below a certain threshold. Still, there will be a point below which, if we knew for certain the impact of the information on price, we would not wish to impose liability—i.e, there is a policy-based materiality threshold. And because the tool that we have to measure that impact—the event study—gives us only probabilistic guidance as to what the actual impact of the information on price was, the issues associated with the tradeoff between Type I and Type II errors are the same as with fraud-on-the-market suits. If the same required confidence level is maintained during periods of economic crisis-induced idiosyncratic risk as during normal times and if event studies are required to establish materiality, Type II errors will rise sharply and many more trades that in fact are based on information sufficiently important to meet the materiality threshold will not be actionable. But again, there are no strong reasons to believe that lowering the required confidence level, with the consequent increase in Type I errors, will increase, rather than decrease, social welfare.

270. SEC v. Berlacher, No. 07-3800, 2010 WL 3566790, at *8 (E.D. Pa. Sept. 13, 2010) (referring to “event study as reliable and the best measure of materiality”); see also Mark L. Mitchell & Jeffrey M. Netter, The Role of Financial Economics in Securities Fraud Cases: Applications at the Securities and Exchange Commission, 49 Bus. Law. 545, 572–84 (1994) (collecting cases).

Ultimately, just as with fraud-on-the-market actions during periods of crisis-induced spikes in volatility, event studies are a comparatively less powerful tool in these circumstances for discriminating between information that is and is not material in insider trading cases despite their advantage of being less subjective. This makes our discussion of alternative forms of evidence concerning proof of materiality in fraud-on-the-market class actions²⁷¹ relevant to insider trading as well. Again, we suspect that increased permissibility of more subjective kinds of evidence during economic-crisis periods is more suitable in SEC enforcement cases because its capacity to use prosecutorial discretion cabins their inappropriate use in a fashion not present in cases brought by private parties.²⁷²

The fact that event studies are, in crisis times, a comparatively less powerful tool for discriminating between information that is and is not material takes on special significance in insider trading cases given our explanations for what causes crisis-induced spikes in idiosyncratic risk. These explanations suggest that more kinds of information that insiders typically possess will be important in crisis times and will have a larger effect on price when they are eventually revealed than they would in normal times.²⁷³ This means that in crisis times, insiders would have many more opportunities to profit from insider trading on nonpublic information that in normal times would not be important enough to warrant deterring by

271. See *supra* section IV.A (examining causation and materiality in fraud-on-the-market shareholder class actions).

272. See *supra* section IV.A. In a nonjury trial involving an SEC Rule 10b-5 action against a defendant accused of insider trading, at least one court has ruled that where the defense has introduced expert testimony based on an event study showing no statistically significant price reaction upon the public release of the information on which the defendant traded and the government, to show materiality, only offered the testimony of a financial expert who did not conduct an event study, the government failed to establish materiality. *Berlacher*, 2010 WL 3566790, at *7. We note that this case did not involve a situation where an economic crisis-induced spike in idiosyncratic risk devalued that defendant's use of an event study.

273. Recall from our earlier examples that the standard error for the typical firm during the recent financial crisis was 3.23%, compared to 1.78% in the average year in normal times. *Supra* text accompanying note 242. This says, very roughly, that average bit information released by such a firm had almost twice the effect on price in crisis as it did in normal times.

privately imposed liability or governmental sanction. Yet assuming that the required level of statistical confidence is not changed, the very same larger standard error means that this larger number of more important bits of inside information that are now important enough to exceed the materiality threshold will on average be no more likely to be found by an event study to be material than they were in normal times when their price impacts were below this threshold.

2. *Materiality Under Section 11 of the Securities Act.* — Section 11(a) of the Securities Act²⁷⁴ imposes liability on the issuer, the underwriters, and certain other persons for misstatements and omissions of required information in a registration statement for a public offering. As with corporate misrepresentations and insider trading under Rule 10b-5, the misstatement or omission is actionable only if it relates to a “material fact.” The doctrinal standard for materiality is the same as under Rule 10b-5.²⁷⁵ Assume that following their issuance, the offered securities trade in an efficient secondary market. If the registration statement contains a misstatement or omission whose subsequent correction changes the security’s price, the doctrinal standard would be met. As a result, here too the courts have relied heavily on event studies in determining whether the misstatement or omission was material.²⁷⁶

Assessing the appropriate role of event studies in determining materiality and damages in section 11 cases and the implications of crisis-induced spikes in idiosyncratic risk requires a little background concerning the system of liability and the statutory measure of damages, including the allocation of burdens of proof and persuasion with respect to these elements.

274. Pub. L. No. 73-22, 48 Stat. 74 (codified as amended at 15 U.S.C. §§ 77a–77aa (2012)).

275. See *supra* note 224 (explaining definition of materiality). The Supreme Court has yet to address the proper standard for materiality under section 11 of the Securities Act, but all the circuit courts that have addressed the question have applied the *Northway* definition of materiality in section 11 actions. See, e.g., *Kronfeld v. Trans World Airlines, Inc.*, 832 F.2d 726, 731 (2d Cir. 1987) (collecting cases).

276. See, e.g., *In re World Access, Inc. Sec. Litig.*, 310 F. Supp. 2d 1281, 1298–99 (N.D. Ga. 2004) (using parties’ event studies in determining materiality).

As for liability, section 11(a) imposes absolute liability on the issuer.²⁷⁷ When an issuer offers equities for sale and makes a price-inflating misstatement or omission, it receives the inflated sales price. Imposing liability simply returns the inflation to the buyer. At least in a transaction-cost-free world (including the absence of judicial error), doing so makes sense whatever the level of the issuer's culpability, since the net effect will be as if the offering was priced correctly in the first place. Imposing absolute liability also creates incentives to take cost-effective steps to avoid such price distorting behavior.²⁷⁸

Now consider damages. For a plaintiff still holding her securities at the time of judgment, section 11(e) provides that the prima facie damages measure is the difference between the price paid for the security and its value on the date of suit, with the difference presumed to be caused by the misstatement or omission.²⁷⁹ But this crude measure and presumption obviously does nothing to correct for market-wide or firm-specific factors other than the misstatement or omission that may have influenced the security's price after the plaintiff's purchase. To address this gap, section 11(e) in effect shifts the burden of proving loss causation from the plaintiff to

277. See, e.g., *Hutchinson v. Deutsche Bank Sec. Inc.*, 647 F.3d 479, 484 (2d Cir. 2011) (applying section 11's scheme of absolute liability for issuers).

278. Section 11(a) also imposes absolute liability on other participants in the process: the issuer's top managers, directors, and underwriters. However, section 11(b) grants these other participants an affirmative defense: They have no liability if they can show that they engaged in adequate due diligence. 15 U.S.C. § 77k(a), (b)(3). The overall liability scheme for these other participants is designed to motivate each of them, particularly the lead underwriter, to independently investigate the issuer and to participate actively in the drafting of the registration statement. Commentary by persons intimately involved with the creation of the Securities Act confirm that this in terrorem arrangement for imposing damages in the absence of adequate investigation was a critical part of the legislative plan to promote full disclosure. See William O. Douglas & George E. Bates, *The Federal Securities Act of 1933*, 43 *Yale L.J.* 171, 173 (1933) (noting penalties are both "compensatory" and "*in terrorem*" in nature and therefore are "set high to guarantee that the risk of their invocation will be effective in assuring that the 'truth about securities' will be told"); Felix Frankfurter, *The Securities Act: II*, *Fortune*, Aug. 1933, at 54, 109 (praising in terrorem effects of Act); see also *Feit v. Leasco Data Processing Equip. Corp.*, 332 F. Supp. 544, 581 (E.D.N.Y. 1971) (explaining courts must hold underwriters to high standard of diligence "since they are supposed to assume an opposing posture with respect to management"); *Escott v. BarChris Constr. Corp.*, 283 F. Supp. 643, 696–97 n.48 (S.D.N.Y. 1968) ("The purpose of Section 11 is to protect investors In order to make the underwriters' participation in this enterprise of any value to the investors, the underwriters must make some reasonable attempt to verify the data submitted to them. They may not rely solely on the company's officers or . . . counsel.").

279. § 11(e), 48 Stat. at 83 ("The suit authorized under subsection (a) may be either (1) to recover the consideration paid for such security with interest thereon, less the amount of any income received thereon, upon the tender of such security, or (2) for damages if the person suing no longer owns the security.").

the defendant by granting the defendant an affirmative defense if it can prove the absence of loss causation; damages are reduced to the extent that defendants can show that events other than the misrepresentation or omission caused the price drop.

The implications of economic-crisis-induced spikes in idiosyncratic volatility on determining materiality and damages/loss causation under section 11 track those under Rule 10b-5. Each side will bear the burden of proving statistical significance with respect to the element on which it bears the burden of proof. The plaintiff, if it seeks to prove materiality through an event study, will succeed at doing so by showing a statistically significant drop in share price when the corrective disclosure was made. In turn, the defendant will use an event study to prove that the drop in security price was caused by events other than the corrective disclosure. The fact that each party will likely contest the other's event study further complicates this event-study duel. Despite the complications, the analysis remains essentially the same. If the same standard of statistical significance is maintained during highly volatile times as in normal ones, Type II errors will rise sharply.²⁸⁰ Suppose the plaintiff is required to establish materiality through use of an event study. Crisis times will substantially diminish section 11's capacity to create the situation that would have prevailed if the offering had been priced correctly in the first place and to deter future price-distorting misstatements and omissions in registration statements.

The crisis-induced increase in idiosyncratic volatility has the same impact on the defendant's effort to show that factors other than the corrective disclosure caused all or part of the price drop. Thus, the higher standard of statistical significance will affect both the plaintiff and the defendant. The relative impact will depend on the facts; the defendant's opportunity to disprove loss causation depends on the presence of other events in the relevant period that may have

280. Again, there are no strong reasons to believe that lowering the standard, with the consequent increase in Type I errors, will increase social welfare.

affected the issuer's price. However, the limits on the capacity of event studies to separate the effects of multiple factors make section 11's shift in the burden of proof to the defendant on this issue significant.

C. Judicial Deference to Board Rejections of Premium Acquisition Offers and Substantive Coercion

To this point, our discussion of the legal implications of the pattern of, and alternative explanations for, crisis-induced spikes in idiosyncratic risk has focused on issues related to the content and their private and public enforcement of the federal securities laws. We turn now to corporate law and in particular to the legal implications of our findings for what has been the most disputed element in state corporate law over the last thirty years: the extent of judicial deference to a board of directors' decision to defend against a hostile takeover.²⁸¹ Over this period, Delaware law evolved to give the board wide discretion when it determines there is a risk of "substantive coercion"—a risk that the shareholders may mistakenly accept a tender offer that the board believes undervalues the corporation. As we will show in this section, the concept of substantive coercion as developed by the Delaware Supreme Court makes little sense in normal times. Here we consider whether there is a circumstance in which the concept might make sense: Can crisis-induced spikes in idiosyncratic risk of the sort discussed in Parts I and II make a substantive coercion claim plausible where it would not be credible in normal times? In

281. See Jack B. Jacobs, Fifty Years of Corporate Law Evolution: A Delaware Judge's Retrospective, 5 Harv. Bus. L. Rev. 141, 154–55 (2015) (arguing rise of hostile takeover litigation contributed to changing legal standards for corporate boards). Justice Jack Jacobs, a recently retired Delaware Supreme Court Justice who sat on the Chancery Court during this period, described the matter as follows:

[The Delaware courts] created an entirely new . . . set of standards [for reviewing boards of director's actions] in the landmark cases of *Unocal*, *Revlon*, and *Blasius* . . . [which] were needed to address new realities and issues arising out of novel legal and financial technologies, in order to solve the problem of whether and how boards should respond to hostile corporate takeovers. That evolution was game-changing. It reshaped the governance of boards and the conduct of all players, including legal and financial advisors, in the area of mergers and acquisitions.

Id.

particular, we examine two takeover cases where an economic crisis figured centrally in both the bidder's decision to make the offer and the target board's decision to oppose it: *Air Products and Chemicals, Inc. v. Airgas, Inc.*²⁸² and *Quickturn Design Systems, Inc. v. Mentor Graphics, Inc.*²⁸³

1. A Brief Review of the Evolution of Delaware Takeover Law. — The emergence of the hostile takeover boom in the 1980s subjected traditional corporate law to the equivalent of a stress test. The largest and most tendentious corporate transactions in history created serious doctrinal cracks, the most important of which was allocating, in the face of a hostile tender offer, final decision rights among directors, shareholders, and courts.²⁸⁴ As the law developed, the breadth of the board's discretion to constrain shareholders from approving a hostile offer came to depend upon the court's assessment of the board's professed belief that the offer presented a "threat" to corporate policy.²⁸⁵ An important element of the potential threat was whether shareholders, even with full information, would mistakenly (in the board's view) tender their shares to a hostile bidder. The threat that fully informed shareholders would make this mistake is termed, awkwardly, "substantive coercion."²⁸⁶

Framing the concept most generously, tolerance of board decisions to block a hostile takeover depends in part on how accurately share prices predict the value of the company in the incumbent management's hands. If these prices are thought to be relatively accurate, it is harder for the target board to justify preventing its shareholders from deciding themselves whether to

282. 16 A.3d 48 (Del. Ch. 2011).

283. 721 A.2d 1281 (Del. 1998).

284. Ronald J. Gilson, *The Fine Art of Judging*: William T. Allen, 22 Del. J. Corp. L. 914, 914 (1997).

285. See, e.g., *Unocal Corp. v. Mesa Petroleum Co.*, 493 A.2d 946, 954–55 (Del. 1985) ("[D]irectors must show that they had reasonable grounds for believing that a danger to corporate policy and effectiveness existed because of another person's stock ownership.").

286. As will become apparent, one of the authors bears at least half the responsibility for this very poorly turned phrase. See *Unitrin, Inc. v. Am. Gen. Corp.*, 651 A.2d 1361, 1384 (Del. 1995) (citing Ronald J. Gilson & Reinier Kraakman, *Delaware's Intermediate Standard for Defensive Tactics: Is There Substance to Proportionality Review?*, 44 Bus. Law. 247, 267 (1989) [hereinafter Gilson & Kraakman, *Intermediate Standard*], as origin of term "substantive coercion").

accept a hostile offer made at a premium over that price. The increase that we have documented in idiosyncratic risk during times of economic distress provides, for the first time, a potentially coherent core to the concept of substantive coercion, a term that has become progressively both more important to Delaware takeover law and more empty of analytic content. Understanding this point, however, requires a short detour along a very long road: the development of Delaware's takeover law.

The modern law of takeovers began with the Delaware Supreme Court's decision in *Unocal v. Mesa Petroleum*.²⁸⁷ There the court resolved the conflict between two contending positions over who could decide whether a hostile takeover would succeed: Should the board be prevented from interfering with the offer so that shareholders decide whether to accept a hostile bid or should the board have the power to prevent shareholders from making that choice?²⁸⁸ The *Unocal* court rejected both contending positions in favor of creating for itself what appeared to be a regulatory role: The court would decide whether the hostile offer presented a threat and and if so, whether the board's response was proportional to the threat identified.²⁸⁹

Following *Unocal*, a law review article appeared that influenced the further evolution of Delaware takeover law: Ronald Gilson and Reinier Kraakman's *Delaware's Intermediate Standard for Defensive Tactics: Is There Substance to Proportionality Review?*.²⁹⁰ Anticipating the possibility that the Delaware Supreme Court might be too sympathetic to a board's claim that

287. 493 A.2d 946.

288. Compare Gilson & Kraakman, Intermediate Standard, supra note 286, at 821, 831–48 (arguing conflict of interest inherent in management defensive tactics “exposes the invalidity of defensive tactics in tender offers and delineates a general principle governing management’s appropriate role in the tender offer process”), with Martin Lipton, Takeover Bids in the Target’s Boardroom, 35 Bus. Law. 101, 115–16 (1979) (“There is no reason to remove the decision on a takeover from the reasonable business judgment of the directors.”).

289. See *Unocal*, 493 A.2d at 955 (“If a defensive measure is to come within the ambit of the business judgment rule, it must be reasonable in relation to the threat posed.”).

290. Gilson & Kraakman, Intermediate Standard, supra note 286. Chancellor Chandler assesses the article’s impact on Delaware takeover law in *Air Prods. & Chems., Inc. v. Airgas, Inc.*, 16 A.3d 48, 96–98 (Del. Ch. 2011) (citing article and noting adoption of its proposed test into Delaware case law).

it knew better than the shareholders, the authors sought to provide the court with a framework for responding to such claims in a way that would cabin what would meet the test. From this effort came the awful term “substantive coercion”: the risk that even in the face of full disclosure, target shareholders still might mistakenly accept a hostile bid that is lower than the company’s fundamental value.²⁹¹ To make a claim of substantive coercion credible, the authors would have required a good deal more than just management’s predictable claim that the market price undervalued the company’s shares. The board also would have to state clearly the source of the mispricing and management’s plans for correcting it.²⁹² At the least, the discipline imposed by re-

291. Gilson & Kraakman, Intermediate Standard, *supra* note 286, at 267.

292. *Id.* at 268 (arguing substantive coercion allegation “requires a coherent statement of management’s expectations about the future value of the company” and “showing of how—and when—management expects a target’s shareholders to do better”). Then—Vice Chancellor (now Chief Justice) Strine highlighted the problem that an unconstrained claim of substantive coercion would present:

As a starting point, it is important to recognize that substantive coercion can be invoked by a corporate board in almost every situation. There is virtually no CEO in America who does not believe that the market is not valuing her company properly. Moreover, one hopes that directors and officers can always say that they know more about the company than the company’s stockholders—after all, they are paid to know more. Thus, the threat that stockholders will be confused or wrongly eschew management’s advice is omnipresent.

....

... Professors Gilson and Kraakman—from whom our courts adopted the term substantive coercion—emphasized the need for close judicial scrutiny of defensive measures supposedly adopted to address that threat[.]

Chesapeake Corp. v. Shore, 771 A.2d 293, 327, 329 (Del. Ch. 2000). Vice Chancellor Strine then quotes Professors Gilson and Kraakman to show what “close judicial scrutiny” would entail:

To support an allegation of substantive coercion, a meaningful proportionality test requires a coherent statement of management’s expectations about the future value of the company. From the perspective of shareholders, substantive coercion is possible only if management plausibly expects to better the terms of a hostile offer—whether by bargaining with the offeror, by securing a competitive bid, or by managing the company better than the market expects. To make such a claim requires more than the standard statement that a target’s board and its advisers believe the hostile offer to be ‘grossly inadequate.’ In particular, demonstrating the existence of a threat of substantive coercion requires a showing of how—and when—management expects a target’s shareholders to do better.

....

The discipline imposed by requiring management to state clearly just *how* it intends to cause the price of the company’s shares to increase is a critical check on knee-jerk resort to assertions that a hostile offer’s price is inadequate. For example, if management believes that the price of a hostile offer is inadequate

quiring this showing would force management to specify the metric by which their performance going forward should be measured if the offer were defeated.

In *Paramount Communications, Inc. v. Time, Inc.*,²⁹³ the Delaware “Supreme Court . . . addressed the concept of substantive coercion head on”²⁹⁴ As that court put it in a subsequent case, the “board of directors had reasonably determined that inadequate value was not the only threat that Paramount’s all cash for all shares offer presented, but was *also* reasonably concerned that the Time stockholders might tender to Paramount in ignorance or based upon a mistaken belief, i.e., yield to substantive coercion.”²⁹⁵

The result, it is fair to say, greatly diminished *Unocal* as a serious restriction on a board’s authority to block a hostile takeover by turning substantive coercion into an assumption rather than a standard of proof. Possibly, the Delaware Supreme Court so sharply diluted the restrictions on proving the presence of substantive coercion because neither the court nor advocates of management discretion could articulate a compelling circumstance when fully informed shareholders reasonably could be expected to make a mistake in accepting a hostile offer. If so, Gilson and Kraakman were too clever by half. In their effort to set the standard of proof high, they created a situation where the Delaware Supreme Court, to give the concept broad application, watered down what had to be proved to essentially nothing.²⁹⁶

because the market undervalues the company[] . . . an acceptable statement of the threat to shareholders would require management to describe the steps that it planned to correct the market’s valuation.

Id. at 329–30 (Del. Ch. 2000) (quoting Gilson & Kraakman, *Defensive Tactics*, supra note 290, at 268–69, 274).

293. 571 A.2d 1140 (Del. 1990).

294. *Airgas*, 16 A.3d at 98 (summarizing Delaware Supreme Court’s movement away from *Unocal* in *Paramount*).

295. *Unitrin, Inc. v. Am. Gen. Corp.*, 651 A.2d 1361, 1384 (Del. 1995).

296. Chancellor Chandler reached just this conclusion in *Airgas*:

Inadequate price has become a form of ‘substantive coercion’ as that concept has been developed by the Delaware Supreme Court in its takeover jurisprudence. That is, the idea that *Airgas*’s stockholders will disbelieve the board’s views on value (or in the case of merger arbitrageurs who may have short-term profit

2. *Substantive Coercion and Crisis-Induced Spikes in Idiosyncratic Risk.* — This is where the impact of crisis-induced spikes in idiosyncratic risk comes in. The occurrence of such a spike may present a realistic situation where substantive coercion could occur. In particular, recall our analysis of the potential that the incidence of insider trading would go up as a result of a crisis: The importance of new information about both the company and about its management go up, which may increase the amount and value of insiders' private information compared to that available to the market.²⁹⁷ This crisis-induced gap between management's information and that available to the market may be difficult to communicate because of the crisis even if, unlike the insider trading situation, management is inclined to close the gap through disclosure. In a period when the continued validity of the market's precrisis valuation model of the company is in question and the range of relevant new information expands precisely because of new competing models, simple disclosure of management's information may not solve the problem. In this circumstance, the information gap may be a measure of substantive coercion: For a period of time, management's knowledge advantage over shareholders concerning facts important for valuing the company is significantly increased. Thus, in a period of crisis when an increase in management's knowledge advantage is signaled by a spike in idiosyncratic risk, there is, in the name of substantive coercion, a case for temporarily giving the board more discretion to impede a takeover.

goals in mind, they may simply ignore the board's recommendations), and so they may mistakenly tender into an inadequately priced offer. Substantive coercion has been clearly recognized by our Supreme Court as a valid threat.

Airgas, 16 A.3d at 57.

297. See *supra* section IV.B.1 (arguing kinds of information typically held by insiders are more important in crisis times than in normal times).

In this regard, consider the facts of two well-known takeover cases: *Quickturn Design Systems, Inc. v. Mentor Graphics Corp.*,²⁹⁸ and *Air Products and Chemicals, Inc. v. Airgas, Inc.*²⁹⁹ Each case took place following a financial crisis. Each involved a sharp decline in the target's share price, which appears to have both prompted the bidder to initiate the hostile offer and provided the target a rationale for resistance. As it turns out, however, in each case, the timing of the crisis-induced increases in idiosyncratic risk do not support an application of the theory of substantive coercion articulated above. This is because the crisis-induced spike in idiosyncratic volatility had already dissipated by the time of the hostile offer.

Quickturn was a technology company with whom Mentor Graphics competed.³⁰⁰ Patent litigation brought by Quickturn had, by 1997, resulted in an injunction barring Mentor Graphics from selling certain products in the United States.³⁰¹ There was an associated damages claim by Quickturn that it said could reach \$225 million.³⁰² These developments led Mentor Graphics to consider a hostile acquisition of Quickturn, motivated in large part by a desire to resolve the dispute by extinguishing the claim through an acquisition of the company holding the patent.³⁰³ There was, however, a problem over price. Although Mentor Graphics's investment banker supported the concept of the acquisition, its view was that Quickturn's stock price, which reached \$15.75 during the first quarter of 1998, was too high to make the acquisition worthwhile.

298. 721 A.2d 1281 (Del. 1998), aff'g 728 A.2d 25 (Del. Ch. 1998).

299. 16 A.3d 48.

300. See *Quickturn*, 721 A.2d at 1283 (describing Mentor and Quickturn as publicly traded companies on NASDAQ market specializing in electronic design technology and emulation technology, respectively).

301. See *id.* at 1284–85 (“In December 1997, the [International Trade Commission] issued a Permanent Exclusion Order prohibiting Mentor from importing, selling, marketing, advertising, or soliciting in the United States . . . any of the emulation products manufactured by [Mentor-acquisition] Meta outside the United States.”).

302. See *id.* at 1285 (“Quickturn is asserting a patent infringement damage claim that, Quickturn contends, is worth approximately \$225 million.”).

303. *Id.* (“If Mentor owned Quickturn, it would also own the patents, and would be in a position to ‘unenforce’ them by seeking to vacate Quickturn’s injunctive orders against Mentor in the patent litigation.”).

Things changed abruptly as a result of the Asian financial crisis. By summer 1998 (the second quarter), Quickturn's stock price had dropped to \$6.³⁰⁴ Mentor Graphics's chairman then concluded that "the market outlook being very weak due to the Asian crisis made [the Quickturn acquisition] a good opportunity."³⁰⁵ Mentor Graphics then commenced a tender offer at an approximately 50% premium to Quickturn's crisis-affected market price, but at more than a 20% discount to the precrisis price. In response, and after the requisite investment banking and legal counsel opinions and board meeting discussion, Quickturn took a set of defensive actions, including a delay in holding a shareholder-requested meeting, an action the Chancery Court ultimately upheld. Quickturn also adopted a "dead hand" poison pill³⁰⁶ that it withdrew after similar devices were invalidated in cases involving other litigants and then a "slow hand" poison pill that the Delaware Supreme Court ultimately held violated Delaware law.³⁰⁷

At least superficially, this case presents a circumstance where a claim of substantive coercion is plausible under the framework developed here. If a crisis-induced spike in idiosyncratic risk is in fact occurring at the time of the hostile offer, the spike would indicate that the private information possessed by target management is likely of greater significance than it is in normal times. The market price of the target's shares, which did not reflect this private information, thus could more significantly diverge from a fully informed price—just the

304. *Id.* at 1284 n.5.

305. *Id.* at 1285 (quoting Mentor Graphics Executive Vice President Gregory Hinckley).

306. A dead-hand poison pill is one that can be redeemed only by the directors who adopted it or by successor directors nominated by directors who adopted it. See, e.g., *id.* at 1289 (describing "continuing director" provision in which only those directors could redeem rights). If the adopting board is replaced, the pill will remain in place for its full, typically ten-year term regardless of if the bidder wins a subsequent proxy contest.

307. See *id.* at 1289, 1292–93 (invalidating Quickturn's "Delayed Redemption Provision," which prevented newly elected board members from redeeming rights within six months if purpose of redemption was to transact with specified "interested persons"). A slow hand pill imposes a period following a change in a majority of the board during which the board cannot redeem the pill. The same device is sometimes referred to as a "no-hand" pill. See *Mentor Graphics Corp. v. Quickturn Design Sys., Inc.*, 728 A.2d 25, 27 & n.2 (Del. Ch. 1998) ("This case involves . . . a 'no hand' poison pill of limited duration and scope Some practitioners of the art have described this iteration as a 'slow hand' poison pill."), *aff'd* 721 A.2d 1281.

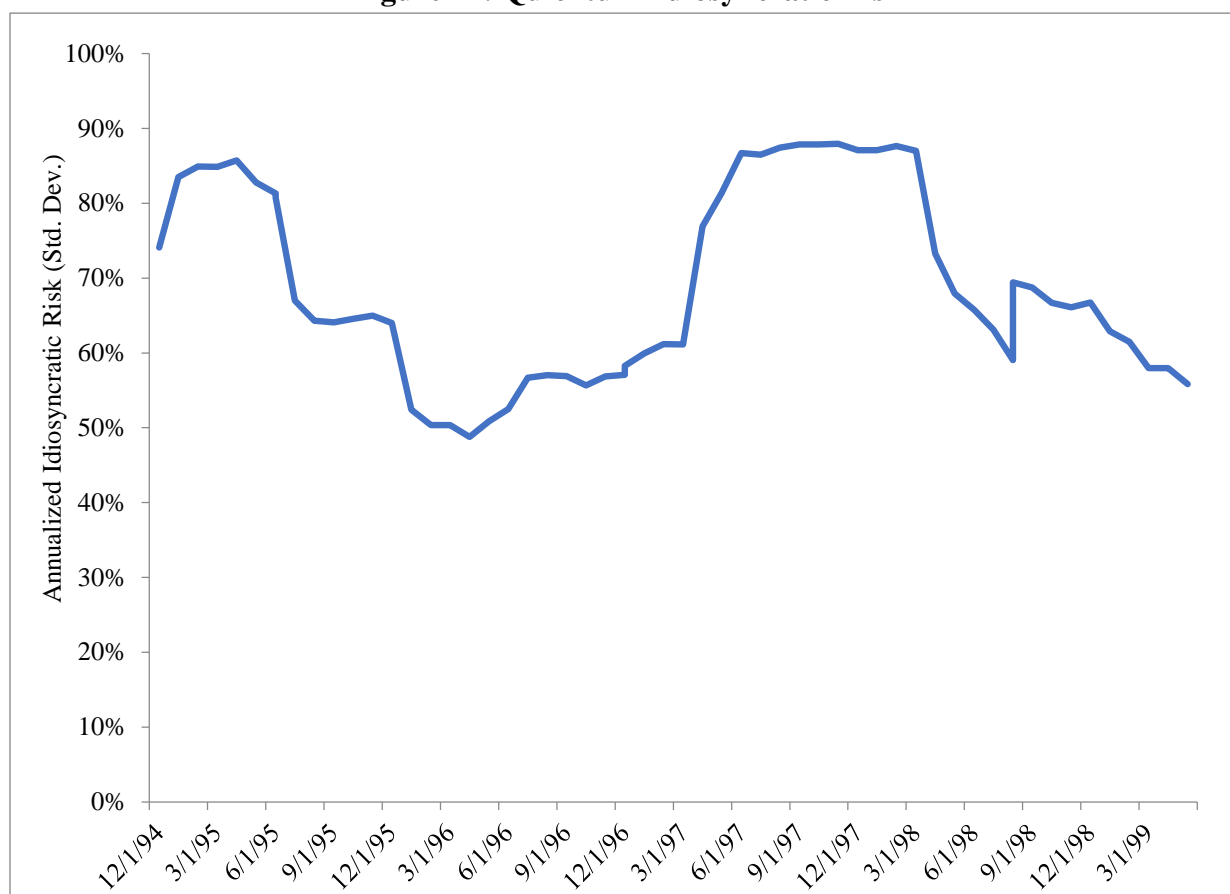
circumstance contemplated by substantive coercion. Target shareholders who compare this market price with the hostile bid could be misled into believing that the bid presented an attractive premium when, judged against management's difficult-to-communicate belief concerning the true value of the shares, it did not. Thus, a spike in idiosyncratic risk could provide evidence of the existence of the kind of situation that Gilson and Kraakman contemplated: one where shareholders could be misled and one where subsequent events—whether share price recovered—could prove whether shareholders were in fact misled.³⁰⁸

The problem with this nice story is that the facts do not support it. Recall from Figure 5 that for the typical issuer during the financial crisis, the pattern of idiosyncratic volatility over time in fact had a spike shape: a rapidly increasing level of idiosyncratic risk followed in approximately one year by a rapid return to precrisis volatility levels. As shown in Figure 14, this is exactly what happened to Quickturn. The company's daily idiosyncratic risk did rise abruptly during the summer of 1997, when the Asian financial crisis surfaced. Although increasingly volatile, its stock price continued a general rise until a peak in the first quarter of 1998, only to collapse following its disappointing second quarter 1998 earnings release.³⁰⁹ Mentor's offer followed this collapse in price. By this time, however, Quickturn's idiosyncratic risk had returned to precrisis levels. Thus, there is no simple link between a crisis-induced spike in idiosyncratic risk and a story that Quickturn's management had a particularly large, difficult-to-communicate, knowledge advantage at the time the offer was made.

308. Gilson & Kraakman, *Defensive Tactics*, supra note 290, at 271.

309. This release reported an 11% drop in revenue from the year-earlier quarter and a quarterly loss of between \$0.12 and \$0.14 per share, compared with a profit of \$0.04 a year earlier.

Figure 14: Quickturn Idiosyncratic Risk



Airgas presents a similar fact pattern, albeit in the context of a different crisis. Prior to the 2008 to 2009 financial crisis, *Airgas* stock traded in the \$40s and \$50s with some periods in the \$60s. With the onset of the financial crisis, the stock dropped—in March 2009 as low as \$27—but recovered to the low \$40s by the time of Air Products’s hostile offer. Just as with *Quickturn*, Air Products had considered a hostile tender offer prior to the crisis, “but did not pursue a transaction at that time because *Airgas*’s stock price was too high. Then the global recession hit, and in the spring or summer of 2009, Air Products’ interest in *Airgas* was reignited.”³¹⁰ Following unsuccessful discussions between the two companies, Air Products then made a

310. *Air Prods. & Chems., Inc. v. Airgas, Inc.*, 16 A.3d 48, 62–63 (Del. Ch. 2011).

hostile offer at \$62, at a time when the Airgas share price was still well below its precrisis level.³¹¹

There followed a series of improvements in the price of the Air Products offer, culminating in a final offer of \$70. A proxy fight led to three Air Product nominees being elected minority Airgas directors, but surprisingly, the Air Products nominees then concluded that the \$70 offer proffered by their nominator was inadequate. In a related case, the Delaware Supreme Court prevented Air Products from accelerating an election that likely would have resulted in Air Products nominees constituting a majority of the Airgas board. The case then came back to the Chancery Court with the central issue being whether the Airgas board could decline to redeem its poison pill, in light of the fact that the pill had “given Airgas more time than any litigated poison pill in Delaware history—enough time to show stockholders four quarters of improving financial results, demonstrating that Airgas is on track to meet its projected goals.”³¹²

Chancellor Chandler made quite clear that in his view substantive coercion should not, as a policy matter, be treated as a valid rationale for declining to redeem the pill when the hostile offer had already been delayed for over a year and there was no serious claim that the shareholders lacked any necessary information. Nonetheless, he felt constrained by Delaware Supreme Court precedent and concluded that Airgas had met the Supreme Court’s empty standard for substantive coercion.

311. The court quotes the CEO of Air Products as stating:

[N]ow is the time to acquire Flashback [the code name for Airgas]—their business has yet to recover, the pricing window is favorable, and our ability (should we so choose) to offer an all-cash deal would be viewed very favorably in this market. To take advantage of the situation, we believe we will have to go public with our intentions.

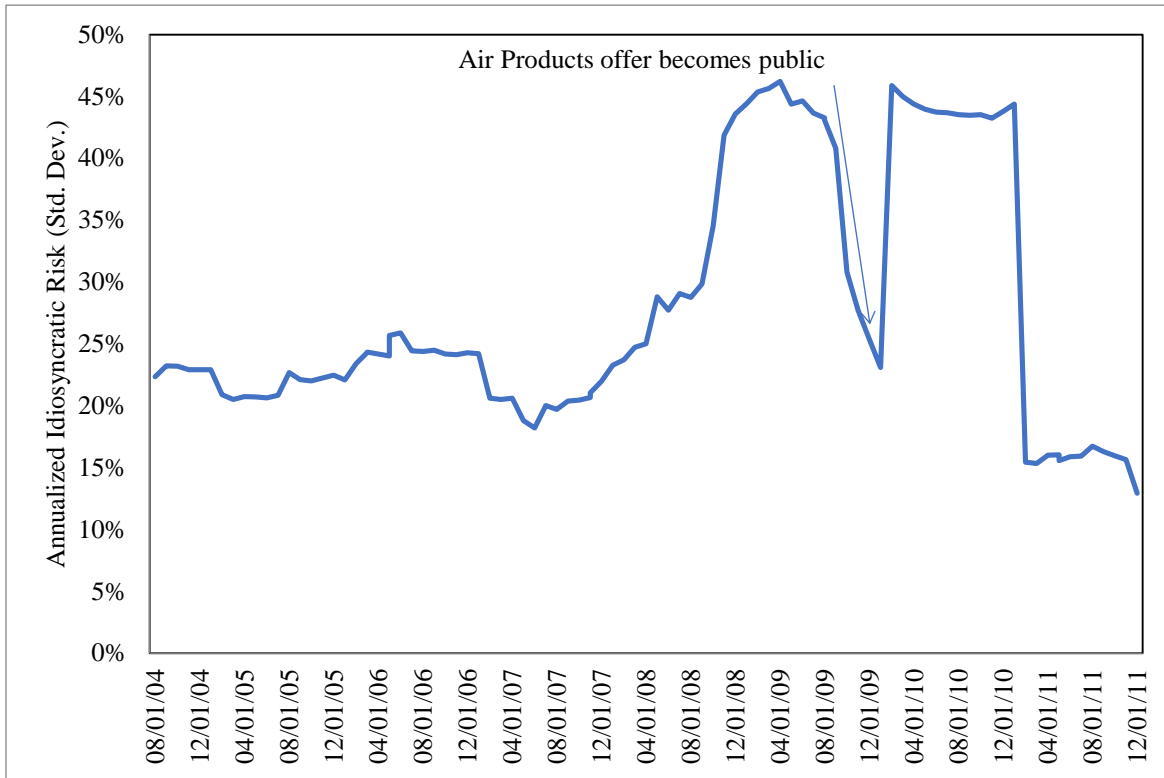
Id. at 68.

312. Id. at 57.

Like *Quickturn*, the timing in *Airgas* makes it difficult to link a crisis-related spike in idiosyncratic volatility to the potential for real substantive coercion. In this respect, the data falsifies the claim that a crisis-induced increase in idiosyncratic volatility might cause shareholders to err in whether to accept the hostile bid because the board possessed an unusually great knowledge advantage over the market. As in *Quickturn*, Figure 15 shows that the increase in Airgas's idiosyncratic risk preceded the Air Products offer, by which time risk levels had returned to normal levels.³¹³ This drop in idiosyncratic risk before the Air Products offer and the further delay in the offer as a result of the litigation, suggests that whatever potential there would have been for a successful substantive coercion claim if the offer had been made earlier, Chancellor Chandler's instincts were right. Such a claim was no longer appropriate by the time of the Air Products offer and certainly not by the time the court issued its opinion.

313. Airgas's idiosyncratic risk increased again *after* Air Products made its offer. This presumably represents uncertainty for a considerable period of time concerning how high Air Products was willing to go and concerning the legal wrangling related to the offer.

Figure 15: Airgas Idiosyncratic Risk



CONCLUSION

Financial economics and associated econometric techniques have come to play a central role in corporate and securities law. This is hardly surprising since at its core, financial economics is concerned with the valuation of financial assets and at their core, corporate and securities law are concerned with establishing rules that facilitate value maximization. Both corporate law and securities law, in turn, depend on the relationship between market prices and value in formulating causes of action. In this chapter, we for the first time document and seek to explain a pattern that existing theory does not predict and existing empirical studies do not reveal: A spike in companies' unsystematic risk has followed every economic crisis from the 1929 stock market crash to the 2008 to 2009 Great Recession. We consider the implications of this pattern for securities law issues, including the proof and measurement of the elements of a cause of action

for securities fraud and insider trading claims and for a central corporate law issue—the extent to which a target board of directors can prevent shareholders from accepting a hostile takeover. More generally we show the workings of what is not widely recognized as a tautology: Law and finance cannot operate independently since one seeks to explain the factors that dictate the value of financial assets and the other seeks to establish rules and institutions that facilitate creating that value.

Chapter 4 Does Capital Bear the Corporate Tax After All? New Evidence from U.S. Corporate Tax Returns

I. Introduction

“We’ll cut business taxes massively. They’re going to start hiring people.”
President Trump during the third 2016 presidential debate.

In both academic and political discussions, the question of who bears the corporate tax remains stubbornly controversial. It is a matter of Republican orthodoxy that the corporate tax falls largely on working Americans by reducing investment at home, leading to fewer jobs and lower wages.³¹⁴ The current administration and Congress seem likely to substantially cut the corporate tax with this as the stated reason. Democrats have been more equivocal, but both major Democratic candidates in 2016 proposed raising additional revenue through the corporate tax. Hillary Clinton proposed raising \$275 billion through “business tax reform” to pay for infrastructure programs, ATLANTIC (Dec. 1, 2015), while Bernie Sanders wanted to raise corporate tax revenue substantially as part of making capital owners pay their “fair share.”³¹⁵

The academic debate too rages on more than fifty years after Arnold Harberger (1962) wrote his seminal article “The Incidence of the Corporation Income Tax.” Harberger showed that under certain assumptions in a closed economy with competitive markets, all capital owners bear the corporate tax burden in the long run. More recent work on open economies, however,

³¹⁴ For example, all of the major Republican candidates in 2016 proposed reducing the corporate rate. *See* Tax Foundation (2016); The same was true in 2012. *See* Tax Foundation (2012).

³¹⁵ In 2012, the Democratic platform called for revenue-neutral corporate tax reform with a reduction in tax preferences paying for a lower statutory rate. *See* Democratic Platform (2012).

offers very different predictions. This research shows that in a small, open economy with perfectly mobile capital and competitive markets, workers must eventually bear the full burden of the corporate tax (see Gordon & Hines (2002)). This is because, unlike capital, workers cannot easily move across countries to escape the tax. Even adjusting for real world imperfections in capital mobility, scholars working in this line have estimated that American workers bear as much as 70% of the corporate tax. The question of the burden of the corporate tax is particularly salient in the United States because the statutory corporate tax rate has remained at the same level for the last 30 years, while tax competition and other factors have caused most other developed countries to significantly cut their corporate tax rates (see Avi-Yonah (2000) on tax competition).

The Treasury Department, which is the principal agency responsible for helping policymakers understand whether tax burdens are distributed fairly,³¹⁶ has also used different models over time. Prior to 2008, the Treasury assigned the entire corporate tax burden to capital owners as a whole. The Treasury tacked in 2008, assigning part of the burden of the tax to labor to reflect changes in recent academic work. The Treasury noted, however, that the tax could have very different effects depending on the source of taxable corporate income. Some corporate income represents the “normal return to capital,” i.e., the return to savers just for the use of their money for a period of time. Taxes on such income may be substantially passed on to American labor as capital flows out of the country to escape the tax. By contrast, under a simple economic model, if a corporation earns “supernormal” returns through monopoly power or some

³¹⁶ The Joint Committee on Taxation (JCT) which also analyzes tax burdens for policymakers has largely thrown up its hands with respect to the corporate tax. In one study in 1993 the JCT assigned the burden of the corporate tax to owners of corporate capital, but otherwise has generally refused to allocated the corporate tax burden due to uncertainty about on whom it falls. This has created odd results, including that the JCT scored the revenue neutral 1986 Tax Reform Act—which decreased personal taxes and increased corporate taxes—as lowering taxes on all income groups (see Auerbach (2005)).

other non-reproducible advantage, taxes on this income are less likely to be passed on to labor. This is because these economic rents are so profitable that investors will not respond to tax increases on this income by shifting capital elsewhere. Even after taxes, these economic rents constitute the investment with the highest return. The Treasury, therefore, adopted a model which assigned 50% of taxes on the normal return to capital on to labor, but left 100% of taxes on supernormal income with the capital owners. (Cronin et al. 2013).³¹⁷

This Article presents new empirical evidence from aggregate U.S. corporate tax return data from 1995-2013 on the composition of corporate taxable income between the normal return to capital and supernormal returns. I show that the corporate tax raised only 4% more money than a hypothetical “R-base” cash flow tax with the same statutory rate. Under this tax, all business expenses can be immediately written off, including long lived assets like buildings and equipment. In addition, all financial transactions are ignored: interest and dividends are not included in revenue, nor is interest deductible as an expense. The 4% figure uses the same assumptions as a number of papers in the literature, but more conservative choices yield an average of about 21% of revenue generated from the normal return.

I then explore the reasons why the tax appears to raise so little revenue from the normal return. Almost certainly part of the explanation is the low risk-free rate, as proxied by the rate on Treasury bonds, during the second half of this period. Nevertheless, even during the mid-1990’s when the risk-free rate was relatively high, only 20-30% of the tax was raised from the normal return, less than one might have otherwise predicted.³¹⁸ Another factor driving down the tax

³¹⁷ As discussed below, the Treasury’s assumption that corporate taxes raised from economic rents falls entirely on the owners of corporate capital is likely too extreme. Nevertheless, I employ this assumption in my analysis below, and one limitation of this chapter is that I reserve for further research the important question of empirically testing the Treasury’s assumption.

³¹⁸ During this period, a tax at the statutory rate (35%) on the nominal risk-free normal return—as proxied by a 10-year Treasury—to the book value of equity in C-corporations would have yielded revenue equal to 70-80% of

raised from the normal return is corporations' increasing investment in intangible capital. This kind of investment is largely given cash flow treatment under the corporate tax. Thus, the tax exempts the normal return on an increasing fraction of corporate investment. Using hand collected tax return data going back to 1957, I find that both decreases the risk-free rate and increases in investment in intangible capital are associated with declines in the portion of the corporate tax estimated to be raised from the normal return.

The similarity of the hypothetical cash flow tax and the existing tax since 1995 has two important implications. First, *in aggregate*, the corporate tax largely exempts the normal return to capital, and the burden instead initially falls on supernormal returns. The primary difference between a cash flow tax and the corporate income tax, is that an income tax is supposed to tax the normal, risk-free, return to capital.³¹⁹ The similarity of the revenue raised by the current tax and the cash flow tax shows that only a small percentage of the current tax is raised from the normal return to capital. Therefore, using the Treasury's model, capital owners bear nearly all the tax. This result differs from previous analyses done using the Treasury model which assign 18-24% of the tax to labor. Thus, while cutting the corporate tax may improve the performance of new and existing corporate equity, if the Treasury model is correct, it is unlikely to substantially increase wages or create jobs.

Second, these results suggest that the U.S. could transition the corporate tax to a cash flow tax and enjoy a simpler and somewhat more efficient tax system without having to substantially raise the statutory rate in order to be revenue neutral. A rents tax would equalize the

actual corporate tax collections. Likewise, a tax on the real risk-free normal rate of return would have raised about 50% of actual corporate tax revenue.

³¹⁹ I put aside the issue for now of whether the cash flow tax would also fall, in part, on the return to risk. To the extent it does, it would do so in the same manner as the current tax. Under the Treasury's model taxes on the return to risk are also borne by capital owners (Cronin et al. 2013). I return to this question in greater detail in Part VII, *infra*.

treatment of debt and equity. This would reduce the use of debt financing and make the economic system more resilient and less prone to financial crises (Schularick & Taylor (2012)). In addition, the average portion of tax raised from the normal return disguises significant differences across industries. For industries which use primarily fixed, physical capital (e.g., mining or utilities), the portion of the corporate tax raised from the normal return can still be significant. By contrast, in industries which use tax-preferred intangible assets, very little tax is raised from the normal return. This leads to tax-favored firms investing in projects with lower pre-tax rates of return than projects rejected by tax-disfavored firms. Correcting these inefficiencies by adopting a tax which uniformly exempts the normal rate of return would improve economic performance and likely raise wages.³²⁰

Nevertheless, the relatively small difference in revenue raised by the hypothetical cash flow tax buttresses contentions that administrative complexity should be a first-order priority when deciding whether to switch to cash flow taxation (Weisbach 2004). For example, a cash flow tax would eliminate the need to track and tabulate depreciation deductions which the Treasury estimates imply costs businesses on the order of \$20 billion annually (Treasury 2011).³²¹ Yet the difficulty of taxing financial services firms under an R-base tax and of distinguishing real transactions from financial transactions suggests that the administrative gains should not be oversold.

Two limitations of this study should be noted. First, it is not clear Treasury's assumption that 100% of taxes on supernormal returns stay with capital owners is justified. In fact, it is likely some of these rents taxes will be shifted in the long run in an open economy (see Devereux

³²⁰ Cutting the corporate tax rate reduces the size of the excess burden and therefore might increase wages, but the lost revenue must eventually be replaced with other distortionary taxes, and therefore the effect on after-tax wages is ambiguous.

³²¹ Treasury estimates that Form 4562 requires 448 million hours annually to comply with.

& Griffith 2003), although less than taxes on the normal rate of return.³²² Because I nevertheless use the Treasury's model, this pushes my results toward the tax falling on capital owners. Second, like past papers, my primary results do not adjust for problems with "loss offsets," which firms receive when they lose money (IRC § 172). These offsets are worth significantly less than their face value because some firms do not generate enough profits to use the offsets to reduce their current and future taxes, and the offsets are neither indexed for inflation nor interest. The insufficiency of loss offsets will make some of the existing tax and a cash flow tax fall on the return to risk, and can make a cash flow tax actually fall partly on the normal return. As a robustness check, I find that my methodology underestimates the percent of the current tax raised from the normal return by up to 10% due to the asymmetry of the tax code.

The remainder of the chapter is organized as follows: Part II gives an introduction to corporate tax incidence analysis and modifications of the traditional models to analyze open economies and imperfect competition; Part III explains why a cash flow tax exempts the normal return, while an income tax does not; Part IV discusses two approximate measures of how much tax is raised from the normal return; Part V describes how the hypothetical rents tax is constructed by replacing depreciation and amortization deductions with an immediate deduction for all new fixed investment and inventory, and by removing net-financial income from the tax base; Part VI presents the primary results showing that the rents tax would have raised only 4% less income than the current tax from 1995 to 2013; Part VII explores some of the implications of the results, and examines the reasons why the cash flow tax and the existing tax raise similar amounts of revenue; Part VIII concludes.

³²² The "incidence" estimated here, as in the Treasury model, is limited to estimating the distribution of the burden of taxes actually levied, and does not include an estimate of the size or distribution of the excess burden created by the corporate tax. Note, however, that the estimates suggest most of the tax is raised from infra-marginal returns/economic rents or the return to risk. As discussed below, taxes raised from these sources should have relatively low excess burden.

II. LITERATURE

Harberger Models

The simplest model of corporate tax incidence concludes that the tax stays where it lands initially: on corporate shareholders in proportion to their ownership.³²³ Taxes frequently induce changes in behavior which shift the tax burden, however, and a more satisfying general equilibrium approach was proposed by Arnold Harberger in 1962. Harberger analyzed a closed economy, with perfectly competitive corporate and non-corporate sectors each producing a different good, and fixed economy-wide amounts of labor and capital which could move freely across sectors. Because there is perfect competition and constant returns to scale, Harberger precludes economic rents: all capital income is the normal return.³²⁴ Intuitively, the tax has two effects in this model. First, corporations switch from using capital to using labor, reducing demand for capital. This places a burden on capital owners by lowering the return to capital. Second, the cost of the good produced by corporations will rise, lowering demand, and shifting capital and labor out of the corporate sector. This shift can also change who bears the tax burden, and its impact depends on a number of parameters.³²⁵ Using reasonable estimates of these parameters, Harberger concluded that capital in fact bore the entire tax in the U.S. over the long term.³²⁶

³²³ As Auerbach (2005) observed, the direct ownership analysis is still useful because changes in the corporate tax often burden the owners of existing corporate capital in ways that cannot be shifted.

³²⁴ In this model, a tax on corporate capital's economic income will cause capital to flow out of corporations into the non-corporate sector. This in turn drives down the return to capital in the non-corporate sector until it is equal to the after-tax return in the corporate sector. Thus, in the long run, the burden on corporate capital owners—if there is any—must be the same as for owners of non-corporate capital since each will earn the same amount after taxes.

³²⁵ Namely the second effect depends on the relative capital-intensity of the corporate and non-corporate sector, the elasticity of substitution between capital and labor in each sector, and the elasticity of demand for the goods produced by the sectors.

³²⁶ The Treasury model is confined to long-term incidence and so I also put aside questions about the tax burden during the transition to the long-run outcome. Readers should note, however, that the transition period can be quite

Countless modifications to Harberger's approach have been made, but the most relevant here are relaxing the assumption of a closed economy and perfect competition. The importance of modeling international trade and capital flows has grown significantly during recent years.³²⁷ If the assumption of no economic rents is retained, a Harberger-style model gives very different predictions when it is extended to a small open economy with perfectly mobile capital. As Roger Gordon and Jim Hines (2002) summarize:

In a small open economy a tax on the return to domestic capital has no effect on the rate of return available to domestic savers since the domestic interest rate is determined by the world capital market. Domestic investment falls in response to higher tax rates. For firms to continue to break even, in spite of the added tax, either output prices must rise or other costs must fall by enough to offset the tax. When output prices are fixed by competition with imports, the tax simply causes the market-clearing wage rate to fall. As a result, the burden of the tax is borne entirely by labor or other fixed domestic factors. Of course, actual conditions differ quite a bit from this model. The U.S. is not a small economy whose policies will only negligibly affect world interest rates and goods prices. Moreover, capital is not perfectly mobile and the goods produced in the U.S. are not perfectly substitutable for those produced abroad.

Scholars have reached quite different conclusions on who bears the burden of a corporate tax on the normal return to capital depending on how they account for these latter complications. For example, Jennifer Gravelle (2013) summarizes four of these studies, with one finding as much as 70% of this tax is borne by labor. She observes that the studies with the largest estimates of labor's burden do not account for at least one of the complexities discussed above. She argues that using the best estimates of limits on capital mobility, international product substitution, and the traditional parameters involved in the Harberger model, all of the studies

important.

³²⁷ For example, from 1987 to 2003 the value (at current cost) of private fixed capital in the U.S. grew at 5.3% per year from \$10.7 trillion to \$24.8 trillion, while U.S. owned assets abroad grew at 11% per year from \$1.4 trillion to \$7.4 trillion and likewise foreign owned U.S. assets grew by 11.5% per year from \$1.4 trillion to \$8.2 trillion (Auerbach (2005)).

would yield *roughly* that 40% of the U.S. corporate tax is borne by labor and 60% by capital.³²⁸

Citing Gravelle's study and others like it, the Treasury model assumes that 50% of the tax on the normal return is borne by American labor. (Cronin et al. 2013).

Taxing Economic Rents

Like adding the effect of international trade, relaxing the assumption of no economic rents appears to be increasingly important in analyzing the U.S. corporate tax burden. In a closed economy, with neoclassical production functions and savings behavior, a tax on these rents is non-distortionary and has traditionally been found to be borne by the owners of corporate capital.³²⁹ Intuitively, investors will still invest in rent earning assets, which even after taxes, earn more than the normal return. Therefore, the amount of capital invested and the allocation of that capital will not change if a tax on rents is introduced or increased. Taxing these rents imposes no excess burden whether they result from declining returns to scale in a competitive market, or from market power held by firms with non-reproducible advantages like brand-names, know-how, or other intellectual property. Following these studies, the Treasury model assumes that 100% of taxes falling on economic rents stick with capital owners.³³⁰

³²⁸ Other international studies have found similar tax burdens or higher burdens on labor, *see, e.g.*, Desai, Foley, and Hines (2007); *but see* Suarez-Serrato & Zidar (2016) in the domestic setting.

³²⁹ Fane (1984), in response to Feldstein (1977), points out that for a tax on rents to be borne entirely by the rentiers, it must be analyzed "follow[ing] 'the traditional practice in incidence analysis' of considering compensated taxes." The Treasury model implicitly employs this assumption and I follow it here.

In reality, of course, it is likely that this assumption does not hold, making a tax on rents accruing to the owners of corporate capital, equivalent to a compensated tax along with a lump sum transfer away from those owners. This transfer can affect relative prices, in turn affecting the ultimate incidence of the tax. Fane observes, however, that in many models "lump-sum redistributions of income do not affect relative prices. Even when they do, the incidence effect is often small relative to the size of the redistribution." Nevertheless, an additional avenue for further research would be to calibrate a model with representative agents to help us understand whether less (or indeed more) than 100% of an uncompensated tax on rents accruing to the owners of corporate capital is borne by those owners.

³³⁰ Even in a closed economy, more complex models sometimes shift the burden of compensated rents taxes off the owner of the rent producing asset or find that a rents tax could be somewhat distortive. For example, "where investors must either commit a large chunk of capital or none at all . . . taxes on pure rents may affect both the composition and level of investment" (Griffith, Hines, and Sørensen (2008)). In addition, in more complicated models of imperfect competition, taxes on rents may also affect the size of the rents extracted (Davidson and Martin

Moving to an open economy complicates the question of who bears taxes on rents. Some economic rents earned by U.S. corporations are closely tied to the U.S. either because of natural resources or other immobile productive factors (e.g., key employees who will not move or agglomeration economies like Silicon Valley). The effect of taxing these rents is covered by the traditional analysis described above. Other rents, however, are firm-specific, like brand names, and may allow the firm to earn supernormal returns wherever it chooses to produce. Taxing these rents may reduce domestic investment as firms move production abroad to escape the rents tax.³³¹ This will lead to American labor bearing some of the tax on these rents, for the same reason as with taxes on the normal return in the open-economy Harberger-style models discussed above.

Estimating the Portion of Corporate Taxes Raised from the Normal Return

There is a small body of literature which estimates how much of the corporate tax is raised from the normal return and how much from supernormal returns. Roger Gordon and Joel Slemrod (1988) analyzed data from 1983 and concluded that a cash flow tax that exempted the normal rate of return would actually have raised more money than the existing tax, and thus that no money was raised that year from the normal return by the corporate tax. Gordon, Laura

(1985); Liu and Altshuler (2013)).

In addition, there may be “rent sharing” which allows labor to capture some economic rents which might otherwise flow to capital owners. For example, high union wages in the auto industry from roughly 1950 to 1980 are usually interpreted as rent sharing. See Alder et al. (2017). In these cases, corporate taxes on economic rents may reduce the rents shared with labor and thus such taxes would fall in part on employees of the firm. This may be true at the very high-end of the wage scale as well, because CEO compensation is likely partly rent-sharing. See Piketty, Saez, and Stantcheva (2014).

³³¹ Although the U.S. has a world-wide system of taxation, it allows companies to defer taxation on the profit of foreign subsidiaries until they are repatriated, effectively lowering the rate on income earned by U.S. corporations in lower tax jurisdictions.

Kalambokidis, and Slemrod (2004a) (“GKS”) performed a similar calculation on data from 1995 and concluded that moving to a cash flow tax would reduce corporate tax revenue by 16%.³³²

In 2013, Julie-Anne Cronin, Emily Y. Lin, Laura Power, and Michael Cooper of the Treasury’s Office of Tax Analysis (OTA)—using somewhat different assumptions from GKS—and data from 1999-2001, 2004, and 2007 concluded that 37% of the tax fell on the normal return. Since then, Treasury has distributed the corporate tax burden using this analysis, assigning 18% ($37\% \cdot 50\%$) of the tax to labor. Most recently, Power and Frerick (2016) used the Cronin et al. methodology to examine 1991-2013 and find that supernormal returns as a portion of taxable income are increasing over the period ranging from 60% at the start of the period to 75% at the end. As discussed in detail below, most of the differences between the GKS results and those of Cronin et al. are attributable to the assumptions used in calculating the changes to the R-base and which summary figure the authors use to approximate the portion of the tax raised from the normal return.

III. CASH FLOW TAXATION AS A TAX ON ECONOMIC RENTS

In simple models, a textbook cash flow tax on businesses initially falls only on economic rents and thus does not distort marginal investments, while a textbook income tax falls both on economic rents and the normal risk-free return to capital. Many readers are familiar with these results. Those readers should feel free to skim this section. I briefly review these results below.

All the models discussed are highly stylized. They provide a background for understanding how constructing a hypothetical cash flow tax can help to separate out how much of the current tax is raised from the normal return to capital. I reserve until later a discussion of

³³² Gordon et al. (2004b) also makes a similar calculation for 2004, but using 2000 data adjusted for changes in profits and investment finding switching to an R-Base would have lowered revenue from non-financial C-corps by \$55 billion. Laura Kalambokidis’ (unpublished 1991) dissertation examines the period from 1975-1986 using somewhat different assumptions.

how additional complexities, most importantly, risk and the absence of fully refundable tax losses, affect my conclusions.

A firm will invest in any project in which the project's (properly discounted) income stream meets or exceeds its cost. For the moment, assume that there is no risk and no taxation. Let us also assume that a project will end in year T and be worthless at that point, and the project costs \$1, all paid in period 0, and $e_1 \dots e_T$ is the stream of income the project produces, and r is the risk-free discount rate. The firm (or representative shareholder) will invest if $\sum_{t=1}^T \frac{e_t}{(1+r)^t} \geq 1$. The lowest value project the firm will be willing to invest in earns $\sum_{t=1}^T \frac{e_t}{(1+r)^t} = 1$. Such a project is worth exactly the value of the opportunity cost of the project, which is lending at the risk-free rate, r . This is the "marginal" project if the firm has many projects to choose from. Note that in a perfectly competitive world with constant returns to scale, all the projects a firm can invest in will be like the marginal project, in which the return on the project is exactly equal to its costs. This means that there are no economic rents.

Moving out of the tax-free world, it has long been understood that under certain conditions a tax on real cash flows does not distort marginal investment decisions (Carey Brown 1948). Such a tax gives an immediate deduction for the full cost of the project in period 0 and no depreciation deductions are available later. Assuming the tax is refundable, or that the firm has other income to be offset by the deduction, the firm's marginal investment decision looks exactly as it did in the world where it ignored taxes. Intuitively, if τ is the tax rate, the government becomes a full partner in the project by providing τ percent of the initial capital investment and

taking τ percent of the earnings in all later periods. This means the firm's rate of return does not change.³³³

More formally, earnings in each period are reduced by $\tau \cdot e_t$, where τ is the tax rate, but the cost of the initial investment is also reduced by τ . Thus after-tax earnings are $\sum_{t=1}^T \frac{e_t}{(1+r)^t} (1 - \tau)$, while the cost of the project is now $(1 - \tau)$. Hence for the marginal project, the firm's calculus is the same after-taxes as without taxes because the after-tax earnings on the marginal project are exactly equal to its after-tax cost: $\sum_{t=1}^T \frac{e_t}{(1+r)^t} (1 - \tau) = (1 - \tau) \leftrightarrow \sum_{t=1}^T \frac{e_t}{(1+r)^t} = 1$. Note that the government raises no revenue from this tax in real terms on marginal projects, which the reader will recall earns, r , the risk-free rate. The government gives a subsidy of τ and over time collects revenue with a present value of τ .³³⁴ Thus the normal, risk-free return to capital is untaxed under an ideal cash flow tax. If the project will earn an economic rent such that $\sum_{t=1}^T \frac{e_t}{(1+r)^t} = 1 + n$, the government will collect $\tau \cdot n$, where $n > 0$ is the value of the economic rent. Government revenue is worth, in present value, $\tau \cdot (1+n)$ and it provides a subsidy of τ . Firms will still invest in all such rent earning projects, however, because after taxes the projects still have a positive net present value.

Unlike a cash flow tax, an income tax requires firms to gradually write off assets which last longer than one year. These long-lived assets include fixed capital goods like equipment and structures. In addition, payments for: training employees, executives engaged in long-term strategy, R&D, development of customer relationships, advertising, and other ways of building

³³³ This result can also be thought of as involving a "gross up" where the immediate deductibility of investment allows a firm to invest τ more than it could originally, and this second investment generates a subsidy of τ^2 and so on, which, when the infinite stream is summed, equals $\frac{1}{(1-\tau)}$. Such a subsidy is assumed to be put into marginal projects since all rent producing projects will have already been fully exploited.

³³⁴ Recall the present value of the earnings stream for the marginal project is 1. The government gets τ percent of those earnings in each period through the tax. The tax revenue thus has a present value of $1 \cdot \tau$.

good will benefit the firm over a period of multiple years and would be deducted over time under an ideal income tax. Nevertheless, these expenses are immediately deductible in full under the existing Code.

Capitalized assets do eventually stop producing income for the firm and thus decline in value over time. We can represent this yearly depreciation by comparing the change in the value of the project in year t relative to its value in the previous year. In symbols: $\text{Depreciation}_t = V_t - V_{t-1}$, where V_t is the project's value in period t .³³⁵ In equilibrium, the value of holding onto the project for another period must match the value of selling the project and investing in the safe asset. Thus $(1 - \tau) \cdot (e_t + V_t - V_{t-1}) = rV_{t-1}$. Note, here r is now the *after tax* risk-free rate. Given that $V_T = 0$, this implies $(1 - \tau) \cdot (e_t + 0 - V_{T-1}) = rV_{T-1}$, which can be rewritten as $V_{T-1} = e_T / (1 + \frac{r}{(1-\tau)})$. Repeating this process backward, we can write the value in any period as $0 \leq t \leq T$ as:

$$V_t = \frac{e_{t+1}}{1 + \frac{r}{(1-\tau)}} + \dots \frac{e_T}{(1 + \frac{r}{(1-\tau)})^{T-t}}$$

Thus the firm will use $\frac{r}{(1-\tau)}$ as its discount rate (or cost of capital) and the marginal project under the income tax must earn on average a return of $\frac{r}{(1-\tau)}$ because the income tax taxes risk-free rate of return.³³⁶ Like a cash flow tax, however, an income tax also taxes rents at a rate of τ per dollar of rents.³³⁷

³³⁵ Ignoring taxes, for any period t , the value of the project can be written as $V_t = \frac{e_{t+1}}{1+r} + \dots \frac{e_T}{(1+r)^{T-t}}$.

³³⁶ Note, however, that r , as the after-tax risk-free rate, may be a function of τ as well. If $\frac{r}{(1-\tau)}$ is unaffected by τ —because r falls exactly as τ rises—then the income tax will not affect investment decisions. Even in this case, however, the government will still raise revenue from the normal return to capital.

³³⁷ For simplicity, the way I have written $(1 - \tau) \cdot (e_t + V_t - V_{t-1}) = rV_{t-1}$ assumes that the value of the rent producing asset in the firm's hands rises from 1 to $1+n$ —where n is the present value of the rent—immediately upon purchase and this increase of n is immediately taxed. Nothing turns on this assumption, however.

In reality, the U.S. corporate tax is a hybrid of an income tax and a cash flow tax because it allows companies to expense a portion or all of the purchase price of long-lived assets, rather than requiring them to fully capitalize the assets. As discussed below, R&D, payments to executives to engage in long-term planning, advertising and other expenses which build good will, and employee training probably exceed 50% of total business investment in long-lived assets, yet are immediately expensed under the Code. In addition, in order to encourage investment, Congress has provided for “bonus depreciation” for 11 of the past 15 years. This has allowed firms to immediately expense 30-100% of the purchase price of qualifying equipment, depending on the year. The 2015 PATH Act extended bonus depreciation through 2019 and to me “bonus” appears to be the new normal. Moreover, even without bonus depreciation, the Code’s standard depreciation system (MACRS) is designed to allow firms to recover depreciation deductions faster than economic depreciation takes place (Margalioth 2007).

Under such a hybrid system, the firm’s discount rate and cost of capital will be $\frac{r}{(1-\sigma\tau)}$, where $(1-\sigma)$ is the fraction of the project initially expensed, and the firm is allowed to take σ percent of the economic depreciation as a deduction thereafter.³³⁸ This has the effect of roughly exempting $(1-\sigma)$ percent of the normal rate of return from tax. Intuitively it can be thought of as allowing $(1-\sigma)$ percent of any investment to be taxed under a cash flow regime, while the remaining σ percent of the project is taxed using an income tax. Rents are still taxed at τ since both types of taxes raise τ per dollar of economic rents. As the fraction initially expensed goes to 1, the tax system converges to the cash flow outcome.

I turn now to empirically estimating how much of the corporate tax is raised from the normal return to capital.

³³⁸ The derivation is more complicated than for a pure income tax and I do not show it here. A full exposition can be found in Auerbach (1983).

IV. HOW MUCH OF THE U.S. CORPORATE TAX IS RAISED FROM THE NORMAL RETURN TO CAPITAL? TWO APPROXIMATIONS

Over the last twenty years both the nominal and real risk-free rate have declined substantially, without a matching decrease in corporate income. Thus, even if the U.S. corporate tax was a textbook income tax, the portion of corporate tax revenue raised from the normal return to capital would have fallen. I illustrate this point using two related methodologies.

First, following William Gentry and Glenn Hubbard, I approximate the portion of U.S. corporate income attributable to the normal return to capital by comparing the return to 10-year government bonds with the returns to equity in public companies. Gentry and Hubbard calculated that over the 1980s stocks returned 16.5% annually after including dividends and capital gains. They assumed firms paid on average an effective tax rate of 25%, and therefore grossed up corporate equity returns by this amount for a total of 22% per year ($=16.5\%/0.75$). Over the same period, they found that (approximately) riskless Treasury bonds had returned 10%. They therefore concluded that about 45% ($10\%/22\%$) of the return to corporate equity was attributable to the normal return and the rest to rents or risk and likewise that about 45% of the current corporate tax was raised from the normal return to capital. The Treasury used Gentry and Hubbard's figures when distributing the corporate tax in the years 2008 to 2011 (Cronin et al. 2013). I update these figures below in Table 7:

Table 7:
Corporate Returns Attributable to the Normal Return to Capital
Using Gentry and Hubbard Method

Period		Annual Nominal Return		Corporate Return
Start	End	Stocks ¹	Risk-Free Rate ²	Attributable to Normal Return
1980	1989	23.47%	10.58%	45%
1995	2013	13.14%	4.47%	34%
1995	2002	14.16%	5.74%	41%
2003	2013	12.41%	3.55%	29%
1965	2015	13.09%	6.51%	50%

¹ Using the Center for Research in Security Prices total return index which is the return on a market-cap weighted average of all stocks trading on the NYSE, NASDAQ, and AMEX, including dividends. The return is then grossed up by 25% to account for corporate taxes paid as in Gentry and Hubbard.

² Defined as based on the 10-year T-bill rate. For each year, the risk-free return is based on the average 10-year T-bill rate during that year. Data are from the St. Louis Federal Reserve (FRED).

From 1995 to 2013, the period covered by this Article, the risk-free rate made up 34% of the rate of return on public corporate equities. This is noticeably less than the 45% seen during the 1980s, which in turn was close to the 50% average figure for the last 50 years. In addition, the real return, rather than the nominal return, is arguably the more relevant figure.³³⁹ During 1995-2013, the real normal return was only 22.5% of real corporate equity returns, compared with 35% from 1980-1989 and 34% from 1965 to 2015.

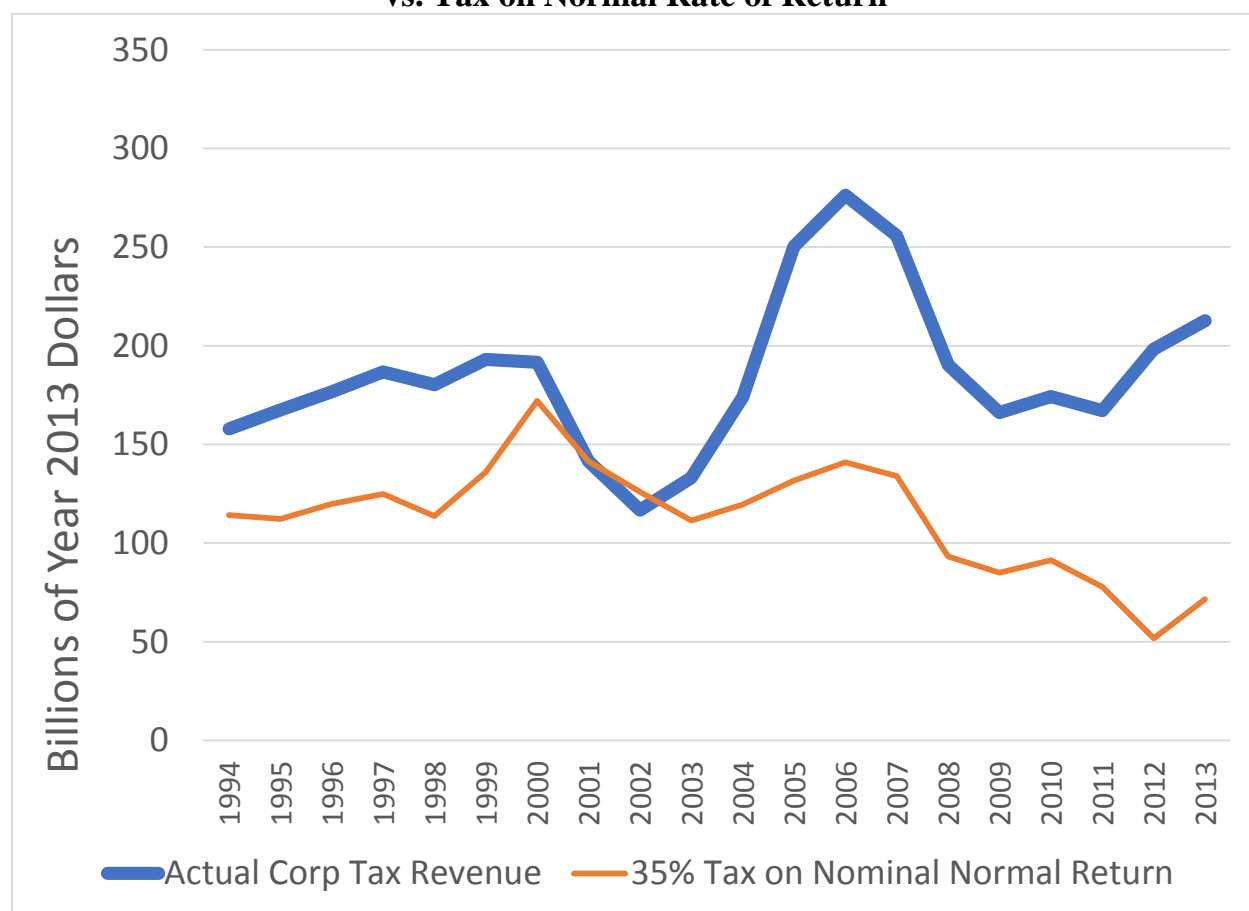
Although Gentry and Hubbard's technique provides a convenient way to approximate the portion of corporate income accounted for by the normal return to capital, I prefer not to interpret it as the percent of corporate taxes raised from the normal return. This is because it does not account for how the corporate tax differs from a textbook income tax with full capitalization of all long-lived assets and no accelerated depreciation.³⁴⁰

³³⁹ An ideal income tax would tax only real income, not nominal income. The corporate tax does, however, tax some nominal gains resulting from inflation by not indexing tax basis for inflation, although this effect is partially offset by the deductibility of nominal interest. See Auerbach (1983).

³⁴⁰ Moreover, this technique is quite sensitive to stock market booms and busts. For example, ending the period of study at the end of 2008 during the financial crisis yields the normal return making up 54% of total corporate

Second, I examine how much an ideal income tax would raise from the normal riskless rate of return for non-financial C-corporations and compare this with actual corporate tax revenue from these firms from 1994 to 2013. I do so by calculating the book value of these C-corporations reported to the IRS and multiplying it by the nominal risk-free rate.³⁴¹ I apply the statutory 35% rate of the current corporate tax. The results are presented in Figure 16:

**Figure 16: Corporate Taxes on Non-Financial C-Corps
vs. Tax on Normal Rate of Return**



During the first half of the period (1994-2003), a 35% tax on the nominal normal return would have made up on average 79% of the actual corporate tax raised. During the second half, it would have composed 48%. The average over the full period was 64%. The decline in the

returns —far more, rather than far less—than the 1980s figure.

³⁴¹ Details of the calculation are discussed in the Appendix, Table 3A-1.

importance of the normal rate of return is even more apparent using the real, instead of the nominal, return. The relevant figures are 46% for the first half and 15% for the second. Although these are approximate calculations, they give us a sense of the upper bound for how much less a cash flow tax might raise than the existing tax, as well helping us to see how the importance of the normal return has declined over the last twenty years.

V. R-BASE CASH FLOW TAX—METHODOLOGY

One way to implement a cash flow tax is to levy a tax on the firm's cash flow in real transactions, while ignoring financial transactions for tax purposes. This is an "R-base" tax as outlined in the Meade Report of 1978. Under such a tax businesses can immediately deduct all real expenses, but cannot deduct interest or dividends paid. On the other hand, firms do not include interest or dividends received in gross income. All of the most recent attempts to measure how much the corporate tax raises from the normal return to capital (Gordon et al., Cronin et al., Power and Frerick) construct an R-base and I do the same. Moreover, the current administration proposal for changes to the corporate tax is an R-base variant with immediate expensing of physical capital for five years and limits on interest deductibility. While an R-base tax has the advantage of simplicity, it provides no easy way to tax financial intermediaries, because most of the revenue that these firms collect for their services is embedded in financial

flows.³⁴² Therefore, like GKS and Cronin et al., I confine my attention to non-financial corporations.³⁴³

I start by applying the same assumptions GKS used, but extending their analysis over the 1995-2013 period using aggregate tax data.³⁴⁴ Although I use the GKS assumptions, I do not use the same summary measure of the portion of the tax raised from the normal return. Instead I use the measure from Cronin et al. to make my results comparable to those used in the Treasury model. This figure measures the change in revenue associated with expensing productive capital (except land) divided by the total taxable income generated by real activities:

$$\frac{Depreciation_t + Amort_t + Depl_t - New\ Invst_t - \Delta Inventory_{t-1,t}}{Taxable\ Income_t - Net\ Fin\ Inc_t}$$

This proxy is adapted from Toder and Reuben (2007), and is designed to roughly account for both debt and equity funded corporate projects, without having to explicitly examine how interest paid by corporations to individuals or partnerships (or financial corporations) is taxed. It can be thought of as an estimate of the portion of the current tax which would be raised from the normal return if all corporate projects were entirely equity funded and corporations had no net

³⁴² Additional problems are raised by the difficulty in separating the embedded interest component in leases, seller-financed transactions, or royalty arrangements. Countries using Value Added Taxes (VATs) face these problems as well and the treatment of leases appears to vary depending on the property in question (equipment, fixed residential property, fixed commercial property, or services/intangibles) and the lease terms (whether the lease is an “operating lease” in which the lessee returns the property to the lessor at the end of the term, or a “finance lease” in which the property will be sold at the end of the term with the proceeds largely going to the lessee). It is not feasible to separate interest embedded in lease and royalty payments in the corporate tax data provided in the Statistics of Income used in this study.

³⁴³ One way to deal with the problems raised by financial firms under an R-base tax is to tax all businesses on real and financial cash flows, known as a R+F base in the terminology of the Meade Report (see AUERBACH, A MODERN CORPORATE TAX (2010) for a proposal to use an R+F base for all businesses). Or, as with the proposed “Growth and Investment Tax Plan,” one could imagine a system where only financial firms would be required to use R+F accounting. See Cunningham and Engler (2012). I do not believe, however, the effect of moving to such a tax can be calculated from public tax data.

In the Appendix Table 3A-2 I include a rough calculation comparing an income tax to an R+F tax by looking at how NIPA R+F cash flows compare with NIPA book profits for non-financial corporations. I find that on average from 1945-2015 an R+F base has been about 20% smaller than book profits, but that relationship is reversed over the last decade with the R+F base being about 20% larger.

³⁴⁴ Both GKS and Cronin et al. methods require that capital grows at the real rate of interest to be accurate.

financial income.³⁴⁵ The downside of using this figure is that the “base” it uses in the denominator may be quite different from the actual tax base. For example, the corporate tax places little or no burden on debt funded projects with individual lenders, but these projects are given equal weight with equity funded projects in this proxy. In Appendix Table 3A-3, therefore, I also present figures based on the estimate of the change in revenue from shifting to the R-base tax divided by the current amount of tax revenue raised, which is the summary figure used in GKS.

To calculate the numerator, I replace depreciation, amortization, and depletion deductions—the corporate tax code’s mechanisms for the gradual cost recovery for capitalized assets—with immediate expensing of all new fixed investment.³⁴⁶ In addition, I allow firms to immediately expense the cost of producing inventory. Under existing law, inventory is capitalized until sold.

In adding these new deductions for immediate expensing and eliminating deductions associated with gradual cost recovery, the methodology makes no distinction between firms with positive net income and “deficit firms” (i.e. those with losses). In other words, giving a deficit firm a new deduction is treated the same way as giving a new deduction of the same size to a firm with net income, and vice-versa for removing existing deductions for gradual cost recovery.

³⁴⁵ Holding fixed behavior, GKS look at both the change in revenue of implementing an R-base corporate tax and from shifting individual income taxation to a consumption tax. Because GKS examine both the corporate and the individual side, they can control for the net effect of not taxing the normal return at either the corporate or individual level. Using the GKS measure, but looking at only the corporate tax will miss tax placed on the normal return to corporate projects through taxing interest paid by corporations to individuals and partnerships.

³⁴⁶ I measure new fixed investment using BEA Table 4.7, which gives estimates for investment in structures and equipment for non-financial corporations. I then adjust this figure to account for the fact that the BEA includes S-Corps in this figure, by multiplying by the ratio of non-financial C-corp depreciation deductions to total depreciation deductions of non-financial C and S-Corps.

I also adjust for differences in coverage between the BEA’s definition of fixed investment and what is depreciable under the Tax Code by using BEA Table 7.13 (e.g. treatment of foreign branches etc.). My measure of investment differs slightly from Gordon et al., who use the Commerce Department’s “Annual Capital Expenditures” survey. BEA Table 4.7 uses data from that survey but also uses other sources.

This can cause the methodology to over or understate the effect switching to expensing would have, holding fixed behavior, on actual tax revenues in a world without full refundability.³⁴⁷

To calculate the denominator, I remove from the existing tax base net financial flows: by (1) removing taxable interest received from the base,³⁴⁸ (2) adding interest deductions back to the base, and (3) removing domestic dividends (net of the dividends received deduction).³⁴⁹

Following Cronin et al., I also remove all foreign dividends, including constructive dividends from controlled foreign corporations from the tax base.³⁵⁰ Finally, I remove all capital gains and losses as well as other gains and losses.³⁵¹

One potential complication—as observed in Slemrod (2007)—is that the method of comparing the actual tax to a hypothetical R-base tax to understand the tax placed on the normal

³⁴⁷ If, on net, new deductions are disproportionately concentrated in deficit firms, then my methodology will likely overstate the cost of switching to expensing in terms of tax revenue. In fact, deficit firms do appear to receive a slightly disproportionate share of new expensing deductions. For example, depreciation makes up 2.6% of deductions of firms with net income, but 3.5% of deductions in deficit firms in 2013. On the other hand, as discussed below, deficit firms present a problem for expensing because imperfect loss offsets can cause even a cash flow tax to fall partially on the normal return to capital.

³⁴⁸ I exclude interest from state and local bonds since that interest is not taxed. See IRC § 103.

³⁴⁹ For domestic dividends, the amount “removed” from the base is smaller than the gross size of domestic dividends, because I adjust for the “dividends received deduction.” See IRC § 243. Companies owning less than 20% of the dividend paying company receive a 70% deduction, those owning more than 20% but less than 80% receive an 80% deduction, and those owning more than 80% of the dividend paying firm receive a 100% deduction. On average over the period this meant that almost exactly 80% of gross domestic dividends were deductible. See SOI Table 20 for years 1995-2013.

³⁵⁰ Only a relatively small tax is collected on repatriated foreign dividends as a result of the accompanying foreign tax credits. However, if the tax base is measured starting from “Income Subject to Tax” as reported in the *Statistics of Income*, this figure includes the gross amount of foreign dividends without any adjustment for accompanying foreign tax credits. Thus, to reach an estimate of domestic C-corp taxable income attributable to real activities, the gross foreign dividends should arguably be removed from “Income Subject to Tax.”

I say arguably only because the “Income Subject to Tax” is understated compared to actual income from real domestic activities because of profit shifting by multinational corporations. The empirical importance of profit shifting out of the U.S. is hotly contested, see Dharmapala (2015), but taking a semi-elasticity of 0.8 with respect to tax rate differentials, profit shifting will have a significant effect on the measured “Income Subject to Tax” net of financial income. See Power and Frerick (2016).

³⁵¹ Under a cash flow tax, the sale of used assets is immediately taxable in full to the seller, but the buyer gets an immediate deduction, with exactly offsetting consequences if both firms are taxable corporations with net income. Thus, expensing all new investment and eliminating all capital and non-capital gains should lead to the same outcome as expensing both new and used assets and including the sale of used assets in the seller’s taxable income. I use the first strategy. Note, however, that this can create problems for sales of used assets into and out of the corporate sector.

return to capital relies on the corporate capital stock growing at the rate of interest. If capital grows slower than the rate of interest, the method will understate the portion of the tax raised from the normal return and vice-versa if corporate capital is growing more quickly than the risk free interest rate. Putting aside operating deductions, recall that deductions under a textbook income tax will be $\delta_t K_{t-1}$ where δ_t is the average economic depreciation of corporate capital in period t , and K_{t-1} is corporate capital in $t-1$. Under the cash flow tax, the relevant deductions are equal to new capital purchased in period t , which is equal to $(\delta_t + \alpha_t)K_{t-1}$ where α_t is the growth rate of corporate capital in period t . For the difference between the cash flow and income taxes to be the tax on the normal return to capital, α_t must equal r_t , where r_t is the normal risk-free return.

In the extreme, we can see the potential for understatement by imagining a text-book corporate income tax in a risk-free world, in which all corporate projects are equity funded, earn the normal return, r , and K_t depreciates at rate δ , but $\alpha_t = 0$ and thus $K_t = K_{t-1}$ because $I_t = \delta K_{t-1}$. In this world, the textbook corporate income tax imposed in period t will raise $\tau \cdot r \cdot K$ and 100% of the tax is raised from the normal return. Applying the GKS method outlined above, however, to period t will yield that 0% of the tax comes from the normal return, because new investment I_t exactly equals economic depreciation allowed by the income tax (δK_{t-1}), and hence the cash flow tax would raise exactly the same amount as the income tax or $\tau \cdot r \cdot K$.³⁵²

One way to deal with this problem is to adjust for the difference between the rate of capital growth based on new capital purchases (net of depreciation) and the risk free rate. In the two-period example, in period t , $I_t = \delta K_{t-1}$ and thus $\alpha = 0$, and the additional deduction under the

³⁵² Under a hybrid income tax, where certain types of capital (e.g., intangible capital) K_1 can be immediately expensed, while other types of capital, K_2 , still require capitalization, it is the growth rate of K_2 which must match the interest rate for the GKS method to be accurate. The relevant correction if α_2 does not equal r , is $(r - \alpha_2)K_{2,t-1}$.

modified cash flow tax would be $(r - \alpha)K_{t-1} = r K_{t-1} = rK$.³⁵³ Thus the adjusted cash flow tax would raise 0, as it should. The difficulty here is in the measurement. Small changes in how either capital growth or the interest rate are measured can lead to very large adjustments once the figure is multiplied by the entire capital stock, indeed enough to swamp the other factors. I present one set of estimates below, but emphasize that it is highly sensitive to the assumptions used. If we are prepared to deal with a reduced sample, another way to deal with this problem is simply to focus on years in which capital growth closely matches the risk-free interest rate.

³⁵³ An equivalent method is suggested in Kalambokidis (1991) in which a corporation prior to the first period of analysis is given full expensing for a deemed “purchase” of all of its capital, and at the end of the last period is taxed on a deemed sale of its capital. In the two-period example above, the cash flow tax will provide full deductions for the capital stock at the end of period $t-1$, at a cost to the Treasury of $\tau \cdot (1+r) \cdot K$ in terms of period t dollars, it will then collect $\tau \cdot r \cdot K$ on cash flows in period t , and τK on the deemed sale of capital at the end of period t . In total the adjusted cash flow tax now correctly raises \$0 in real terms.

VI. RESULTS

Figure 17: Portion of Corporate Tax Lost Switching to Cash Flow Taxation 1995-2013

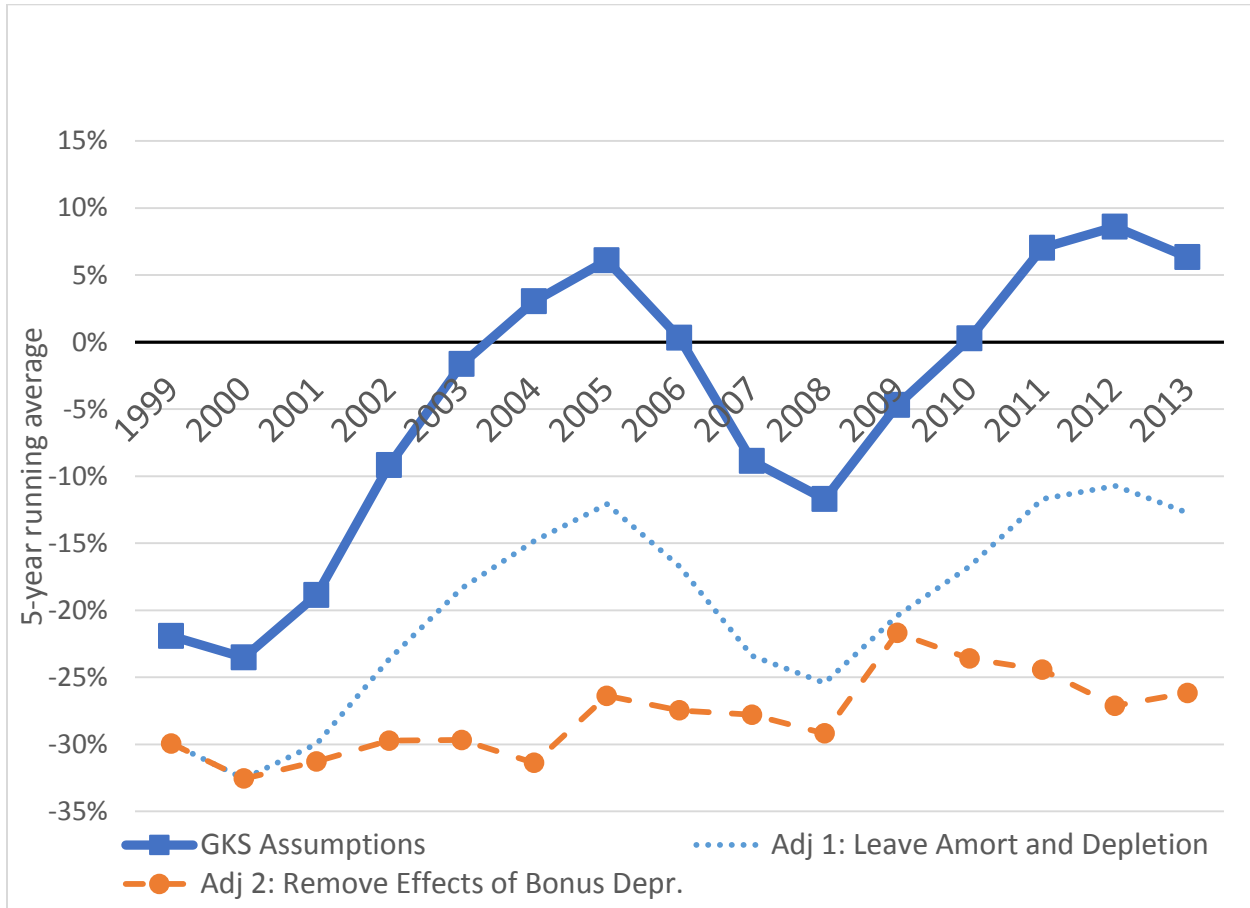


Figure 17 displays the main results. Using the GKS assumptions (the blue solid line with square markers), the average portion of the corporate tax raised from the normal return is 4% over the 1995-2013 period. The periods during which the line is positive indicates that the estimated cash flow tax would raise more money than the existing tax. This can happen during poor economic times when businesses run down inventory (generating larger inventory deductions than immediate expensing of inventory) and bonus depreciation reduces or eliminates the difference between cash flow and actual tax treatment of new equipment investment.³⁵⁴ The other two lines

³⁵⁴ I use a five-year running average to make the patterns on the graph easier to view, but this decision makes poor economic condition take a number of years to fully materialize on the figure.

represent cumulative steps in moving toward the Cronin et al. assumptions. Each line further down the chart includes the previous adjustments.

Adjustment 1, (the dotted light blue line) does not remove depletion and amortization deductions in moving to the R-base. The average portion of the tax raised from the normal return under this assumption is 21%. This roughly follows what Cronin et al. do: they retain depletion for taxing natural resource extraction and make only small changes for expensing what are now amortizable assets.³⁵⁵ GKS by contrast assume that no additional adjustment must be made to the investment figures to accurately proxy investment in new depletable or amortizable assets. There are advantages to each approach.

Intangible assets typically need only be capitalized and recovered through amortization when they are purchased from third parties (Kahng 2014). Thus, these assets can be thought of like sales of “used” physical capital assets as discussed above in note 351, and dealt with by eliminating all capital gains. In addition, amortization includes some physical assets —e.g., IRC § 169 (pollution control devices), § 1400I (structures in revitalization neighborhoods)—which are covered in the investment data. In these cases, eliminating the associated amortization deductions is the appropriate treatment. Likewise, for depletion, depletable exploration costs are already covered in the investment data. Moreover, firms which use percentage depletion may take a deduction which exceeds their basis in the property and thus can be more favorable than expensing.³⁵⁶ These arguments favor removing amortization and depletion in calculating the shift to the R-base as GKS do. On the other hand, the Cronin approach is more conservative,

³⁵⁵ Cronin et al. expense assets amortizable under § 197 (acquired intangibles) and software.

³⁵⁶ The Treasury estimates that “excessive” percentage depletion was a roughly \$1.6 billion tax expenditure for corporations in 2013 (Treasury 2013) out of a total of \$8.9 billion of would-be tax revenue which is shielded by depletion deductions (\$25.4 billion of deductions * .35). This suggests that depletion is more favorable than expensing if 5% is the interest rate and the average resource is fully depleted within 8 years.

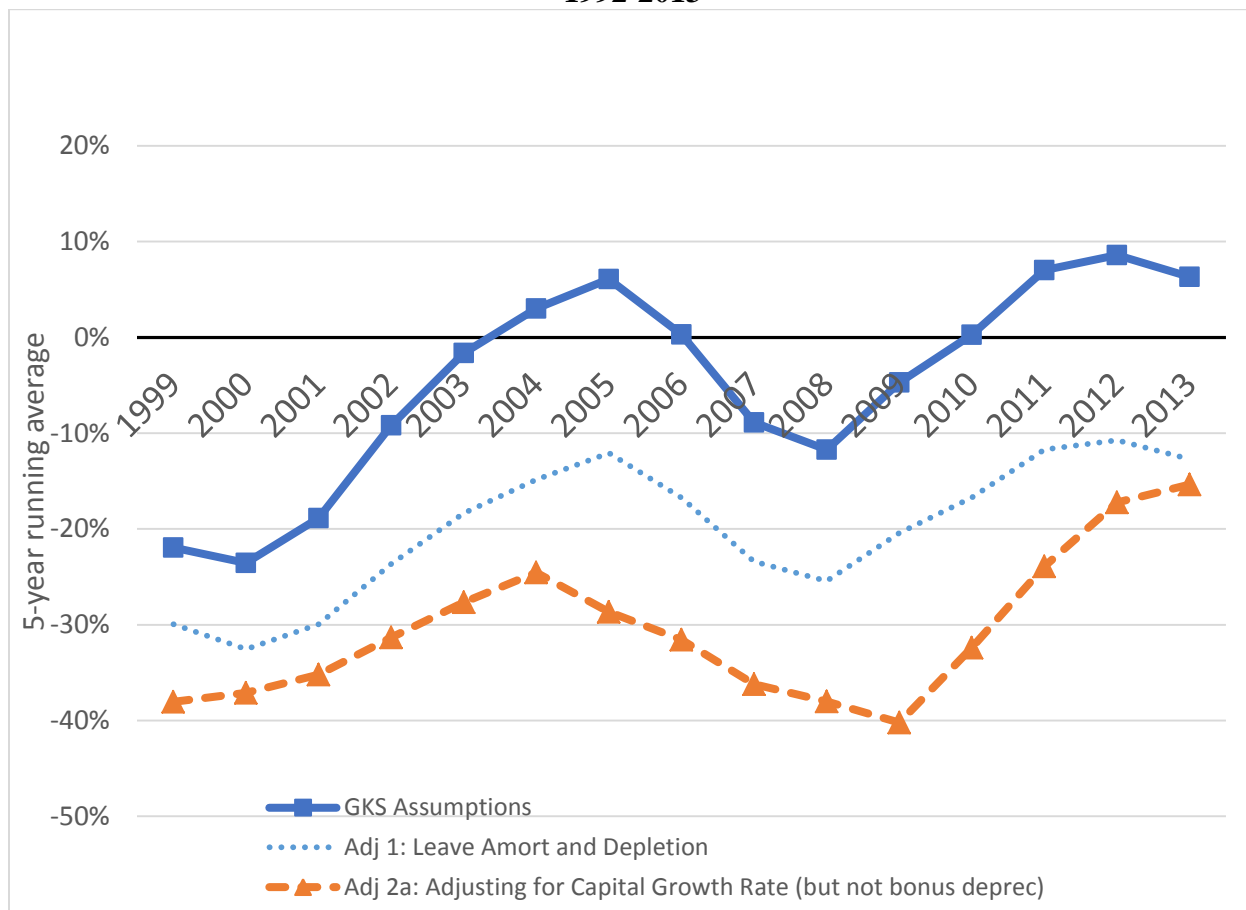
largely leaving amortization and depletion in place where the initial investment associated with those deductions cannot be calculated reliably enough from the tax data.

Second, Cronin et al. strip out the effect of “bonus depreciation.” I have roughly made this calculation in Adjustment 2 (the dashed orange line with circular markers).³⁵⁷ Reasonable minds can differ here, but I believe not trying to remove the effects of bonus depreciation is the better choice. By allowing the immediate expensing of a large portion of the purchase price of many long-lived physical assets, bonus depreciation makes the current tax much closer to a cash flow tax. Moreover, it has been in place for 11 of the last 15 years and therefore it appears to me to be something more like a permanent feature of the Tax Code that we must account for.

Figure 18 presents an analysis adjusting for differences between the rate of growth of physical capital (investment in structures, equipment, and inventories net of depreciation) compared to the risk-free rate proxied by the 5-year constant maturity Treasury rate.

³⁵⁷ I assume that the average MACRS depreciation length for all bonus depreciation eligible equipment is 7 years as in Cronin et al. 2013.

**Figure 18: Portion of Corporate Tax Lost Switching to Cash Flow Taxation
1992-2013**



When measured using these assumptions, adjusting for capital growth increases the portion of the tax raised from the normal return, particularly in the earlier part of the analysis. Again, I emphasize, however, that this pattern is partly dependent on which measures are chosen. Using the theoretically equivalent method outlined in note 353, the portion of the tax raised from the normal return is reduced in nearly all periods and is negative (the adjusted cash flow tax raises more revenue than the existing tax) for most of the 2000s. Likewise, if shorter maturity Treasury notes should be used to approximate the risk free rate—since 5-year notes will likely build in some inflation risk—adjusting for capital growth reduces the portion of the tax raised from the normal return from the baseline.

VII. EXPLANATION AND COMMENTARY

In this section I briefly explain again the two important conclusions that can be drawn from my analysis: (1) that under the Treasury's model, the whole corporate tax burden falls on capital owners, and (2) that the U.S. could transition to a simpler, more efficient, cash flow tax without sacrificing revenue. I then discuss whether risk premiums and imperfect loss offsets could drive the similarity of the revenue of the cash flow tax and the current tax. I argue that this is unlikely for risk. Controlling for imperfect loss offsets does make a noticeable difference. I believe, however, that the primary explanation for the similarity of the two taxes comes from a low real normal return combined with the fact that the Tax Code allows many long-lived expenses like R&D, employee training, advertising, etc. to be immediately written off.

Although corporate taxes make up a relatively small share of the federal budget (11% in 2015), changes to the corporate tax often make up an important part of revenue legislation, as with the Tax Reform Act of 1986, among others (Auerbach 2005). The results above show that under the Treasury's model, the full corporate tax over the last 20 years should be assigned to capital owners. Even under more conservative assumptions (Adjustment 2), only 10% of the tax should be assigned to labor. This differs from previous analyses used by the Treasury and is important for evaluating the distributional effects of cutting the corporate tax which seems likely to occur in the next year. As noted in the introduction, if the Treasury's model is right this cut may increase the performance of new and existing corporate equities, but it will not spark a substantial uptick in domestic hiring or wages.

The evidence also suggests that we could move to a cash flow tax focused on economic rents without having to significantly increase statutory rates or give up revenue. It has generally been assumed that a cash flow tax would raise less money and therefore require higher tax rates

to be revenue neutral (see Auerbach 2010). Raising rates is politically unpopular because it is difficult to explain to the public that the higher rate is balanced by a narrower base. Moreover, as discussed above, some economic rents are firm-specific and taxing these rents at a high rate will drive investment out of the U.S.³⁵⁸ Thus, it is important that the benefits of cash flow taxation are available without needing to raise the statutory rate or even to eliminate various politically thorny corporate tax expenditures like the domestic production activities deduction.

Switching to a cash flow tax would eliminate significant “excess burdens” imposed by the current tax. First, different industrial sectors are taxed unevenly depending on the mix of assets they use in production. Indeed, sectors which use relatively little fixed capital (structures and equipment) and instead rely on human capital and intellectual property may actually find debt financed investment subsidized by the Tax Code on the margin.³⁵⁹ These problems would be eliminated under an R-base cash flow tax. Moreover, an R-base cash flow tax would eliminate the Tax Code’s preference for debt over equity. This would lead to less leverage in the economy which would make it more resilient to financial crises. Last, cash flow taxes are widely regarded as simpler because they do not require complex depreciation schedules and the concomitant accounting by corporations. As noted in the introduction the Treasury estimates that businesses spend over 450 million hours tracking and calculating depreciation.

³⁵⁸ Although it is beyond the scope of this chapter, taxing such rents using a destination basis rather than keeping the traditional source basis for the cash flow tax would likely alleviate this problem. See, for example, the 2016 House Republican plan for international tax reform, which was subsequently abandoned. Switching to a destination basis, however, comes with its own set of concerns: it is untested anywhere in the world, there are questions about its WTO compliance, and it will create large swings in the U.S. dollar (*see* Cunningham and Engler (2012)).

³⁵⁹ These businesses get essentially both the immediate expensing of long lived assets and deductibility of interest. For projects earning the normal return, in order to recover the original investment subsidy provided by immediate expensing, the government must collect $\tau\%$ of future income from the project. By allowing for interest deductibility on top of immediate expensing the government will collect less than $\tau\%$ of the earnings. We should note, however, that the government should recover something like the missing interest deduction by the inclusion of interest income by the lender.

The administrative case for an R-base cash flow tax is not so straightforward, however. Complex as the current code is, it is a devil we know. Implementing a variant of an R-base tax along the lines of the administration plan would require Congress and the Treasury to face a new set of problems. Many common transactions like leases involve both a real and an implicit financial component and separating them would not be easy. Moreover, it can be quite difficult to tell “real” and “financial” transactions apart where the taxpayers find it advantageous to disguise one as the other. Trade credit for example can be embedded in the sale price, making it deductible under an R-base absent clear legislation or regulations. Likewise, a loan could be disguised as a sale of (actually worthless) intangible property combined with a set of royalty payments back to the purported seller (actually the lender). Thus, the importance of administrative considerations in light of the small revenue differences can cut both ways for whether we adopt a cash flow tax.

I move on now to discussing why the cash flow tax raises almost as much as the existing tax:

A. Risk. I have held off until this point discussing the breakdown of what the Treasury defines as “supernormal” returns between the return to risk and economic rents. In a practical sense, it does not matter much to distributing the corporate tax under the Treasury model: taxes on both risk and economic rents are both assigned to capital owners.³⁶⁰ Moreover, taxes on risk premia may raise money in expected value terms, but they likely impose relatively little burden (Gordon 1985). This is because taxes act like insurance by reducing the riskiness of investments. Through taxes the investor gives up a portion of his risk premium and in return he gets a lower

³⁶⁰ The assignment of the tax burden is to different capital owners, however. Taxes on rents are assumed to be borne by owners of corporate equities, while taxes on risk are assumed to be borne by all capital owners. Regardless, the distributional consequences of this distinction are small.

after-tax risk. Giving up part of the risk premium and getting lower risk is basically an even tradeoff for him.³⁶¹ Thus these additional revenues, “have positive expected value but have little market value to the investors who forgo them because of their risk.” (Auerbach, 2005). Indeed, in simple models with perfect loss offsets, income and cash flow taxes impose no burden at all on the returns to risk at all because the investor can “undo” the tax by just investing increasing his investment in risky assets until his after-tax portfolio matches his pre-tax portfolio. (Domar & Musgrave 1944).³⁶²

There are at least two ways we might think about estimating how much role taxes on risk play in the corporate tax base. The first is to use a representative agent with average risk aversion, and ask how much this agent would value stream of revenue produced by the corporate tax? I calculate that the certainty equivalent value of corporate tax receipts from 1995-2013 assuming all the agents in the economy have the same, constant relative risk aversion coefficient of 3. On average the corporate tax raised \$185 billion in 2013 dollars, with a standard deviation of \$39 billion.³⁶³ Risk averse agents would value this income stream the same as a certain

³⁶¹ To make this more concrete imagine a cash flow tax with $\tau = 50\%$. Assume an investor has a normal level of risk aversion and is indifferent between a safe return of 1% or a risky return with an expected return of 3%, where half the time the asset loses 1% and half the time it gains 7% and the investor has \$100. Assume also the investor uses the investment subsidy of \$100 to invest more in the asset he purchased originally.

If he chooses the safe asset he will invest \$200 in the safe asset in period 1, he will get \$202 in period 2 and pay a tax of \$101, leaving him with a 1% return. Likewise, the government’s portfolio is safe: it’s going to get \$101 in period 2 no matter what.

If he invests \$200 in the risky asset, he has a 50% chance of getting \$214 and a 50% chance of having \$198 in period 2 before taxes. After taxes the investor has a 50% chance of having \$99 and 50% chance of having \$107, The government likewise has a 50% chance of collecting \$99 and a 50% chance of collecting \$107, or in expectation \$103. So the government will collect more money in expectation if the investor chooses the risky asset, but these extra revenues have no market value because they come with risks that exactly counterbalance their higher expected value.

³⁶² While investors facing such a tax can “undo” the effect of tax on the risk premium simply by increasing their investment in risky assets, under an income tax they cannot avoid a tax on risk free return of their entire portfolio. (see Warren (1996)).

³⁶³ This analysis assumes that all agents have a utility function which obeys constant relative risk aversion with a coefficient of 3—the most accepted estimates of this coefficient are between 1 and 3 (Gandelman & Rubén Hernández-Murillo (2015)).

payment of \$170 billion, only 8% less, suggesting risk is not driving the large overlap between the actual corporate tax base and the calculated hypothetical cash flow base.

By contrast, taxes on the returns to risk explain much more of the corporate tax base if we extrapolate from the market risk premium. Over time, the equity premium in the U.S. has been about 4.15% (Ayres and Nalebuff (2013)). Under a cash flow tax, the government in essence becomes an equity partner and the tax base will have the same riskiness as the cash flows of the average (public) company.³⁶⁴ From 1995-2001, the average nominal risk free rate, as proxied by 5 year Treasury notes, was about 6% and the real rate was 3.5%, close to the estimated risk premium. Thus, taxes on the returns to risk over this period should have about equaled the tax on the normal return, estimated above at about 20-30% of corporate tax revenue. Moreover since 2002, taxes on risk should have made up a significantly larger part of the corporate tax base (assuming the equity premium is time-invariant).

These very different figures depending on which method we use to think about risk are partly a result of incomplete loss offsets, discussed below, which make the actual tax base less risky than the hypothetical cash flow base. In addition, much of the difference is driven by the fact that the historical equity premium in the United States cannot be explained by ordinary levels of risk aversion. Mehra and Prescott (1985).

B. Loss Offsets. Incomplete loss offsets are another potential explanation for the similarity of the cash flow tax to the current tax. If the government does not fully refund tax losses, then even a cash flow tax can fall on the normal return. Therefore, my estimate of the revenue raised by the R-base cash flow tax may include some tax on the normal return which I do not account for above. Recall that under the cash flow tax the government is supposed to

³⁶⁴ Public companies still dominate the corporate tax base after weighting by size (Auerbach 2005).

provide τ % of the initial capital, and collect τ % of the income from the project. When firms with tax losses for a period invest in new projects they do not receive τ % of the initial capital from the government, but rather additional loss offsets with a face value of τ %. Economists at the Treasury have estimated that firms collect only about 50% of the face value of these loss offsets (Cooper and Knittel 2010). As a sensitivity check, I therefore scale up actual loss offsets, which were used in a given year by 100%. Doing so should roughly approximate the effect of having perfect loss offsets. I estimate that the R-base cash flow tax with perfect loss offsets would raise 10% less than the other estimates with imperfect loss offsets.³⁶⁵ This means that up to 10% of the existing tax may fall on the normal return but not be picked up as falling on the normal return under the GKS and Cronin et al. methods.

C. Low Normal Rates of Return and Expensing of Long Lived Assets.

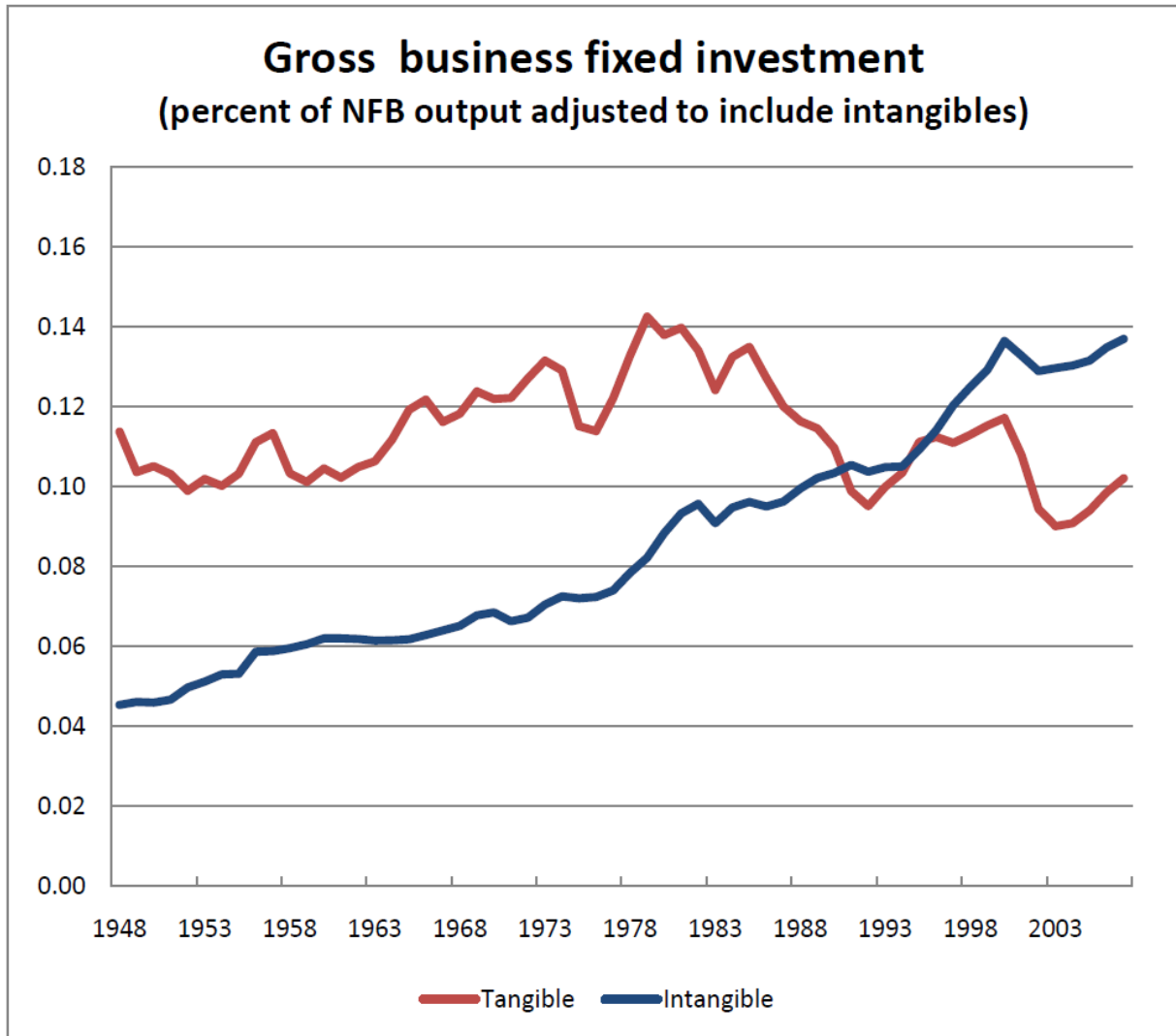
In some ways the finding that we are currently raising relatively little tax from the normal rate of return is obvious. One can lose sight of the fact that the difference between a cash flow tax and an income tax is just a question of timing: when do you recover your basis. In a low interest rate and low inflation environment, the timing does not matter all that much (see Listokin (2016)). For an asset with a 7-year life span, if inflation is 2% and the real interest rate is 2%, then even under straight line depreciation, the difference between expensing and an income tax is 10% of the value of the asset. This is an important wedge, but much smaller than if inflation is 5% and real interest rates are 5%—there the wedge would be 25% of the value of the asset. Yet even during the period during the 1990s when the real risk free rate was 4%, only 20-30% of the tax was raised from the normal return. As noted in the introduction, a tax on the nominal normal

³⁶⁵ This is a large change but even here my baseline result would be that labor bears less than 10% of the tax under the Treasury model. Moreover, previous analyses have not adjusted for imperfect loss offsets and would also find a higher percent of the tax was raised from the normal return if they did so.

rate of return on the book value of C-corps would have raised more than double what the actual tax raised from the normal return.

I believe a substantial factor explaining why corporate tax raised a small percentage of its revenue from the normal return is the tax system's failure to keep up with the rise of intangible property. Corrado and Hulten (2010) estimate that intangible investment broadly construed to include 1) software, 2) innovative property (scientific and non-scientific R&D), and 3) brand investments and organizational investments (e.g., advertising, training of employees, strategic planning by executives) already by the 1990s formed a majority of business investment.

Figure 19:



Source: Corrado and Hulten (2010).

Yet the vast majority of these investments can be immediately expensed as self-developed intangibles. Having given cash flow treatment to more than 50% of investment, we should not be surprised that we collect relatively little from the normal return even when the normal return is high.

I begin to test the importance of changes in the risk-free rate of return, as measured by Treasury notes, and increases in the importance of intangible investment by extending my GKS-

style comparison of the actual corporate tax and a hypothetical R-base tax back to 1957. The collected data are presented in the Appendix in Table 3A-3. Confining my analysis to years in which the growth rate of physical capital is within 200 basis points of the risk-free, rate as proxied by the 5-year constant Treasury rate, I find that decreases in the risk-free rate and increases in intangible capital (as a percent of total capital) are associated with the corporate tax raising a smaller portion from the normal return:

Table 8:
Association between Portion of the Corporate Tax Raised from Normal Return and Risk-Free Rate, Intangible Capital

	Estimated % of Tax Raised from Normal Return
Intangible Assets as a Percent of Total Capital	-0.5221 (0.2526)*
5 Year Constant Maturity Treasury Rate	1.6786 (0.7090)**
R^2	0.33
N	24

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

These results are merely suggestive, but given the strength of the theoretical relationship between these variables and the portion of the corporate tax raised from the normal return, it is quite likely that both factors have played an important role in the decrease in the tax raised from the normal return over time. If these estimates were accurate, the increase in intangible assets from 15% of total fixed assets in 1957 to 37% in 2013 would be associated with a decline of about 10 percentage points in the portion of the tax raised from the normal return. Likewise, the decrease in the nominal risk-free rate from 6% in the mid-1990s to about 1.5% from 2009 to 2013 would have been associated with a decline of about 7.5 percentage points in the portion of the tax raised from the normal return.

VIII. CONCLUSION

I present new empirical evidence from U.S. corporate tax returns which shows that the corporate tax raised very little revenue from the normal return to capital from 1995 to 2013. This means that under the Treasury's model, the full corporate tax burden should be assigned to capital owners, again assuming that taxes on supernormal returns stick fully with capital owners. The results also imply that it would be less costly than commonly thought to move to a simpler and modestly more efficient cash flow tax on corporations.

References

- Alder, Simeon, David Lagakos, and Lee Ohanian, 2017. "Labor Market Conflict and the Decline of the Rust Belt." Working Paper, <https://sites.google.com/site/davidlagakos/home/research>
- Auerbach, Alan J., 2010. "A Modern Corporate Tax." Center for American Progress and The Hamilton Project, Washington, DC.
- Auerbach, Alan J., 2006. "Who Bears the Corporate Tax? A Review of What We Know." *Tax Policy and the Economy* 20, 1-40.
- Auerbach, Alan J., 1983. "Taxation, Corporate Financial Policy and the Cost of Capital." *Journal of Economic Literature* 21(3), 905-940.
- Avi-Yonah, Reuven S., 2000. "Globalization, Tax competition, and the Fiscal Crisis of the Welfare State," *Harvard Law Review* 113(7), 1573-1676.
- Ayres, Ian, and Barry Nalebuff, 2013. "Diversification across Time." *The Journal of Portfolio Management* 39(2), 73-86.
- Russell Berman, 2015. "Hillary Clinton's Modest Infrastructure Proposal." *The Atlantic*, <http://www.theatlantic.com/politics/archive/2015/12/hillary-clintons-modest-infrastructure-proposal/418068/>
- Brown, E. Cary, 1948. "Business-Income Taxation and Investment Incentives." In *Income, Employment and Public Policy: Essays in Honor of Alvin E. Hansen*.
- Cooper, Michael G., and Matthew J. Knittel, 2010. "The Implications of Tax Asymmetry for US Corporations." *National Tax Journal* 63(1), 33-62.
- Corrado, Carol A., and Charles R. Hulten, 2010. "How Do You Measure a 'Technological Revolution'?" *The American Economic Review* 100(2), 99-104.
- Cronin, Julie Anne, Emily Y. Lin, Laura Power, and Michael Cooper., 2013. "Distributing the Corporate Income Tax: Revised US Treasury Methodology." *National Tax Journal* 66(1), 239-262.
- Cunningham, Noël B., and Mitchell L. Engler, 2012. "Prescription for Corporate Income Tax Reform: A Corporate Consumption Tax." *Tax Law Review* 66, 445-486.
- Davidson, Carl, and Lawrence W. Martin (1985). "General Equilibrium Tax Incidence under Imperfect Competition: A Quantity-setting Supergame Analysis." *Journal of Political Economy* 93(6), 1212-1223.

Desai, Mihir A., C. Fritz Foley, and James R. Hines, Jr., 2007, "Labor and Capital Shares of the Corporate Tax Burden: International Evidence." Working Paper, <https://www.aeaweb.org/conference/2011/retrieve.php?pdfid=326>

Dharmapala, Dhammika, 2014. "What Do We Know About Base Erosion and Profit Shifting? A Review of the Empirical Literature." *Fiscal Studies* 35(4), 421-448.

Devereux, Michael P., and Rachel Griffith, 2003, "Evaluating Tax Policy for Location Decisions." *International Tax and Public Finance* 10(2), 107-126.

Domar, Evsey D., and Richard A. Musgrave, 1944. "Proportional Income Taxation and Risk-Taking." *Quarterly Journal of Economics* 58(3), 388-422.

Fane, George, 1984. "The Incidence of a Tax on Pure Rent: The Old Reason for the Old Answer." *Journal of Political Economy* 92(2), 329-333.

Feldstein, Martin, 1977. "The Surprising Incidence of a Tax on Pure Rent: a New Answer to an Old Question." *Journal of Political Economy* 85(2), 349-360.

Gravelle, Jennifer, 2013. "Corporate Tax Incidence: Review of General Equilibrium Estimates and Analysis." *National Tax Journal* 66(1), 185-214.

Gentry, William M., and R. Glenn Hubbard, 1997. "Distributional Implications of Introducing a Broad-Based Consumption Tax." *Tax Policy and the Economy* 11, 1-47.

Gandelman, Néstor, and Rubén Hernández-Murillo, 2015. "Risk Aversion at the Country Level." *Federal Reserve Bank of St. Louis Review* 97(1), 53-66.

Gordon, Roger H., 1985. "Taxation of Corporate Capital Income: Tax Revenues versus Tax Distortions." *The Quarterly Journal of Economics* 100(1), 1-27.

Gordon, Roger H., & James R. Hines, Jr., 2002. "International Taxation." in *Handbook of Public Economics* 4, 1935-1995.

Gordon, Roger H., Laura Kalambokidis, and Joel Slemrod, 2004a. "Do We Now Collect Any Revenue from Taxing Capital Income?" *Journal of Public Economics* 88 (1), 981-1009.

Gordon, Roger H., Laura Kalambokidis, J. Rohaly, and Joel Slemrod, 2004b. "Toward a Consumption Tax, and Beyond." *The American Economic Review* 94(2), 161-165.

Gordon, Roger H., and Joel Slemrod, 1988. "Do We Collect Any Revenue from Taxing Capital Income?" *Tax Policy and the Economy* 2, 89-130.

Griffith, Rachel, James R. Hines, Jr., and P. B. Sørensen, 2010. "International Capital Taxation." in *Dimensions of Tax Design*, 914-996.

Harberger, Arnold C., 1962. "The Incidence of the Corporation Income Tax." *Journal of Political Economy* 70(3), 215-240.

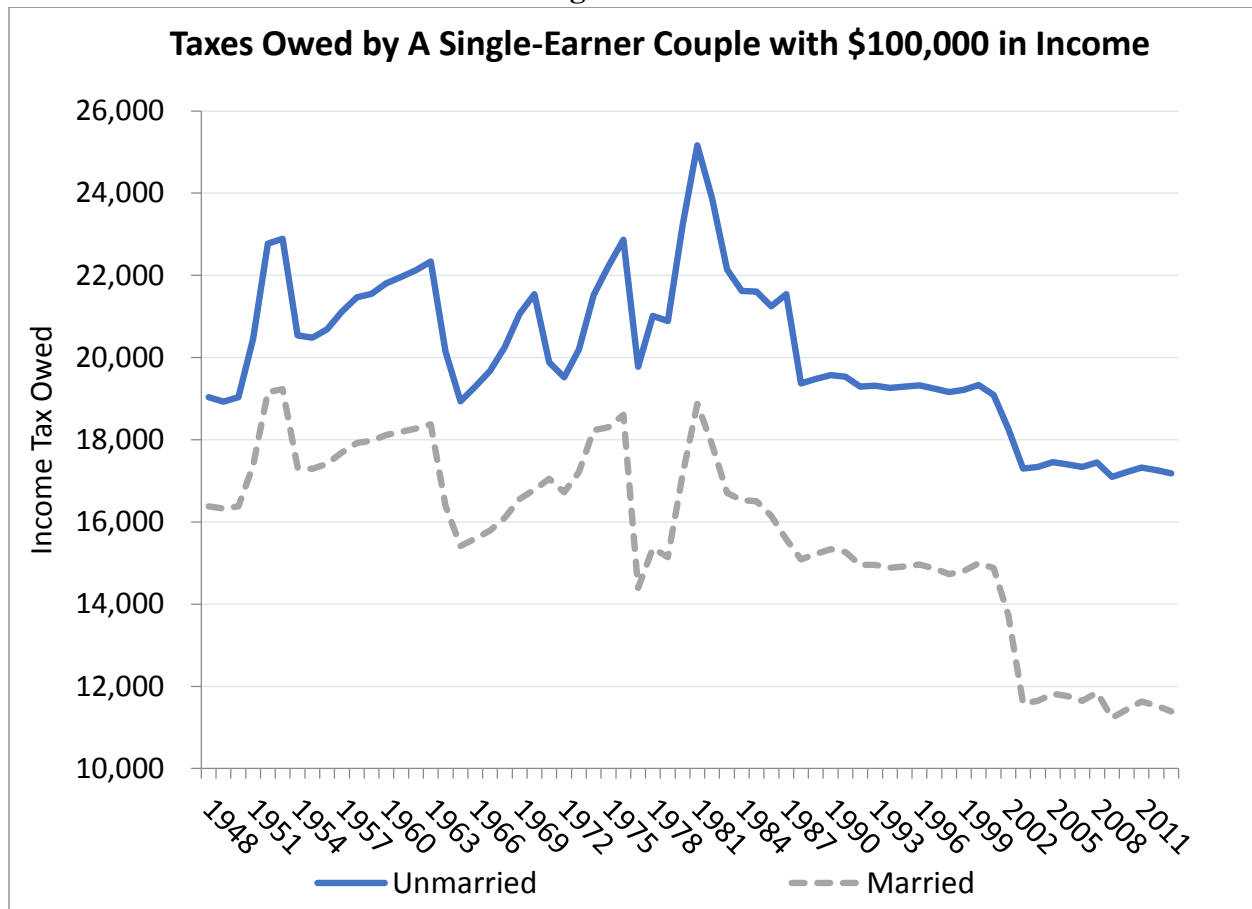
- Kahng, Lily, 2014. “The Taxation of Intellectual Capital.” *Florida Law Review* 66, 2229-2278.
- Kalambokidis, Laura (1991). “What Is Being Taxed? A Test for the Existence of Excess Profit in the Corporate Income Tax Base” (unpublished dissertation).
- Listokin, Yair., 2016, “How to Think About and Teach Income Tax When Interest Rates are Zero.” *Tax Notes* 115(1), 1-4.
- Liu, Li, and Rosanne Altshuler, 2013. “Measuring the Burden of the Corporate Income Tax under Imperfect Competition.” *National Tax Journal* 66(1), 215-238.
- Margalioth, Yoram, 2006. “Not a Panacea for Economic Growth: The Case of Accelerated Depreciation.” *Virginia Tax Review* 26, 493-518.
- Mehra, Rajnish, and Edward C. Prescott, 1985. The Equity Premium: A Puzzle. *Journal of Monetary Economics* 15(2), 145-161.
- Peters, Ryan H., and Lucian A. Taylor, 2017. “Intangible Capital and the Investment-Q Relation.” *Journal of Financial Economics* 123(2), 251-272.
- Power, Laura, and Austin Frerick, 2016. “Have Excess Returns to Corporations Been Increasing Over Time?” *National Tax Journal* 69(4), 831-846.
- Slemrod, Joel, 2007. “Does the United States Tax Capital Income?” In H. Aaron, L. Burman, and C. Steuerle, (eds.) *Taxing Capital Income*
- Schularick, Moritz, and Alan M. Taylor, 2012. “Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870–2008.” *The American Economic Review* 102(2), 1029-1061.
- Stiglitz, Joseph E., 1969, “The Effects of Income, Wealth and Capital Gains Taxation on Risk-Taking.” *The Quarterly Journal of Economics* 82, 262-83
- Tax Foundation, 2016. “Comparing the 2016 Presidential Tax Reform Proposals.” <http://taxfoundation.org/comparing-2016-presidential-tax-reform-proposals>
- Tax Foundation, 2012. “Candidates '12 Tax Plan Comparison.” <http://taxfoundation.org/comparing-2016-presidential-tax-reform-proposals>.
- Treasury Department, 2011. “Supporting Statement [to OIRA under Paper Work Reduction Act].” <https://www.reginfo.gov/public/do/DownloadDocument?objectID=44105901>
- Treasury Department, 2013. “Fiscal Year 2013 Tax Expenditures.” <https://www.treasury.gov/resource-center/tax-policy/Documents/Tax-Expenditures-FY2013.pdf>
- Warren Jr., Alvin C., (1996). “How much Capital Income taxed under an Income Tax is exempt under a Cash Flow Tax.” *Tax Law Review* 52, 1-16.

Weisbach, David A., 2004. "The (Non) Taxation of Risk." *Tax Law Review* 58, 1-58.

Appendices

Appendix 1: Appendix to Do Taxes Affect Marriage?

Figure 1A- 1:



Source: Author's calculations based on historical rates, personal exemptions and standard deduction. Inflation data based on the Consumer Price Index. The single-earner prior to marriage and the married couple are assumed to take the standard deduction. The non-working spouse's personal exemption is assumed to be used before marriage.

**Table 1A- 1:
Additional Results**

	Controlling for Migration		Sub-Regional Comparison
	Excluding Men Moving States After Study Period ¹	Excluding Men Not Residing in Birth-State ²	TX, LA, AL, MS, OK, TN, and AR Only
Post'48*High Income	-0.40% (0.38)	-0.06% (0.40)	-0.88% (0.54)
Post'48*Common Law	-0.09% (0.36)	-0.62% (0.24)**	-0.67% (0.45)
Post'48*Common Law*High Income	0.76% (0.41)*	0.69% (0.45)	1.19% (0.0081)
High Mobilization*Post'48			
R^2 (excluding fixed effects)	0.03	0.03	0.03
N	592,944	403,391	103,452

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Estimated year, state-age-income fixed effects and controls for economic conditions not reported, standard errors clustered at the state level

¹ Excluding all persons who have moved across state lines in 1955-1960. The 1960 census asked respondents where they lived 5 years before.

² Excluding all men who did not live in the same state in which they were born in 1960.

Appendix 2: Appendix to Economic Crises and the Integration of Law and Finance (with Merritt Fox and Ronald Gilson)

The purpose of this Appendix is to describe in more detail how we arrived at our baseline results that idiosyncratic risk spikes during economic crises, and explain additional tests we performed to ensure that our results are robust.

We first walk through the data sources and our main results, and how they compare to Campbell et al.'s work.³⁶⁶ Next, we consider whether industry specific trends can explain spikes in idiosyncratic risk during crises, and find that they do not. Then, we examine whether our results change if we use Fama and French's three-factor model and conclude that this has little impact. We also test whether the idiosyncratic volatility spikes could be caused by crises revealing incompetent management—i.e., the tide going out to reveal the naked swimmers. We find the data inconsistent with what we would expect if volatility spikes were driven by the tide going out. We then analyze whether our volatility spikes are attributable to increases in leverage during crises. We show that increases in leverage explained about 20% of the increase in idiosyncratic volatility during the most recent crisis. Finally, we test whether β instability or serial correlation in idiosyncratic volatility are driving our result. Again, we conclude that these factors do not drive our results.

³⁶⁶ See John Y. Campbell, Martin Lettau, Burton G. Malkiel & Yexiao Xu, *Have Individual Stocks Become More Volatile: An Empirical Exploration of Idiosyncratic Risk*, 56 J. FIN. 1 (2001).

Data

CRSP

Unless otherwise noted, we use data on all firms in the CRSP database. This include all securities traded on the NYSE (from 1925), AMEX (from 1962), and NASDAQ (from 1972) and NYSE ARCA (from 2006).^{367,368}

S&P 100

For some estimates we use the firms in the S&P 100 as of March 2, 2009. There is nothing special about this date; it was the date when the project started. It does, however, introduce a survivorship bias. The firms that were not removed from the index or were added to replace firms that collapsed during the crisis have upwardly biased performances. Overall the firms in the sample had their market caps decline, peak to trough, by about 50%. The market as a whole (as measured by the market cap of all firms in the CRSP database) was down close to 55% peak to trough. Thus, the survivorship bias should not be too severe, and we have no reason to expect that a positive survivorship bias would upwardly bias idiosyncratic risk.

Baseline Estimate

For each firm i in the CRSP data, we ran the “standard” market model regression using Ordinary Least Squares (OLS)

$$Return_{i,t} = \alpha + \beta_i * Return_{mkt,t} + \varepsilon \quad \textbf{Equation 1}$$

³⁶⁷ In addition to common equity shares for firms incorporated in the US, this includes:

(a) common shares of closed end funds, real estate investment trusts, and companies incorporated outside the United States

(b) ADR's and shares of beneficial interest.

Excluding the types of securities listed in (a) and (b) has virtually no effect on the results below.

³⁶⁸ Some smaller firms are occasionally missing shares outstanding information. Missing data is interpolated from the last valid shares outstanding information. This interpolation makes little difference since the aggregate market capitalization of these firms is relatively small.

where the market is the CRSP value weighted index. We ran this model for each firm once a month with an estimation period looking backward roughly one trading year (250 obs).³⁶⁹

Therefore in our figures, the idiosyncratic risk on, for example, March 28, 2001 represents the risk from March 31, 2000 to March 28, 2001 and each regression period will have an overlap of 230 out of 250 observations with the previous one. To illustrate with Abbott Laboratories:

Table 2A- 1
Example of Estimated Idiosyncratic Risk

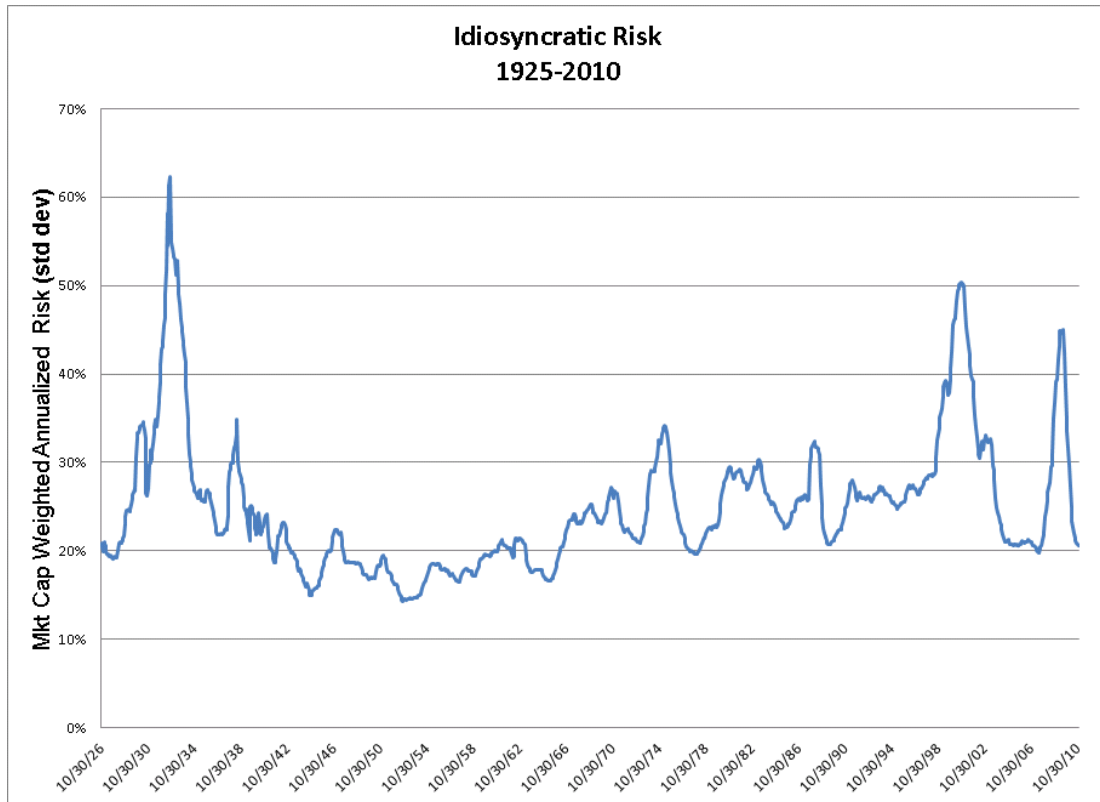
Start Date	End Date	Estimated σ
3/31/00	3/28/01	0.0231
5/1/00	4/26/01	0.0216
5/30/00	5/24/01	0.0212
6/27/00	6/22/01	0.0208
7/26/00	7/23/01	0.0208
8/23/00	8/20/01	0.0203
9/21/00	9/24/01	0.0203

Our estimated σ_i , which is the standard deviation³⁷⁰ of our estimated regression errors ϵ_i , is interpreted as the idiosyncratic risk of firm_{*i*} over the regression period. Each firm's estimated σ is combined to give a market-cap-weighted average. If a firm makes up 5% of the total market capitalization of the sample, its estimated risk will make up 5% of the average.

³⁶⁹ At first glance it might appear that using the Seemingly Unrelated Regressions (SUR) technique would deliver greater efficiency here. However, because the right hand side variable (the market return) is the same across all the firms, OLS and SUR are equivalent.

³⁷⁰ In some figures in our main paper (e.g. Figure 5) we use the variance instead of standard deviation, this is simply σ_i^2 or the standard deviation squared.

Figure 2A- 1
Baseline Estimate

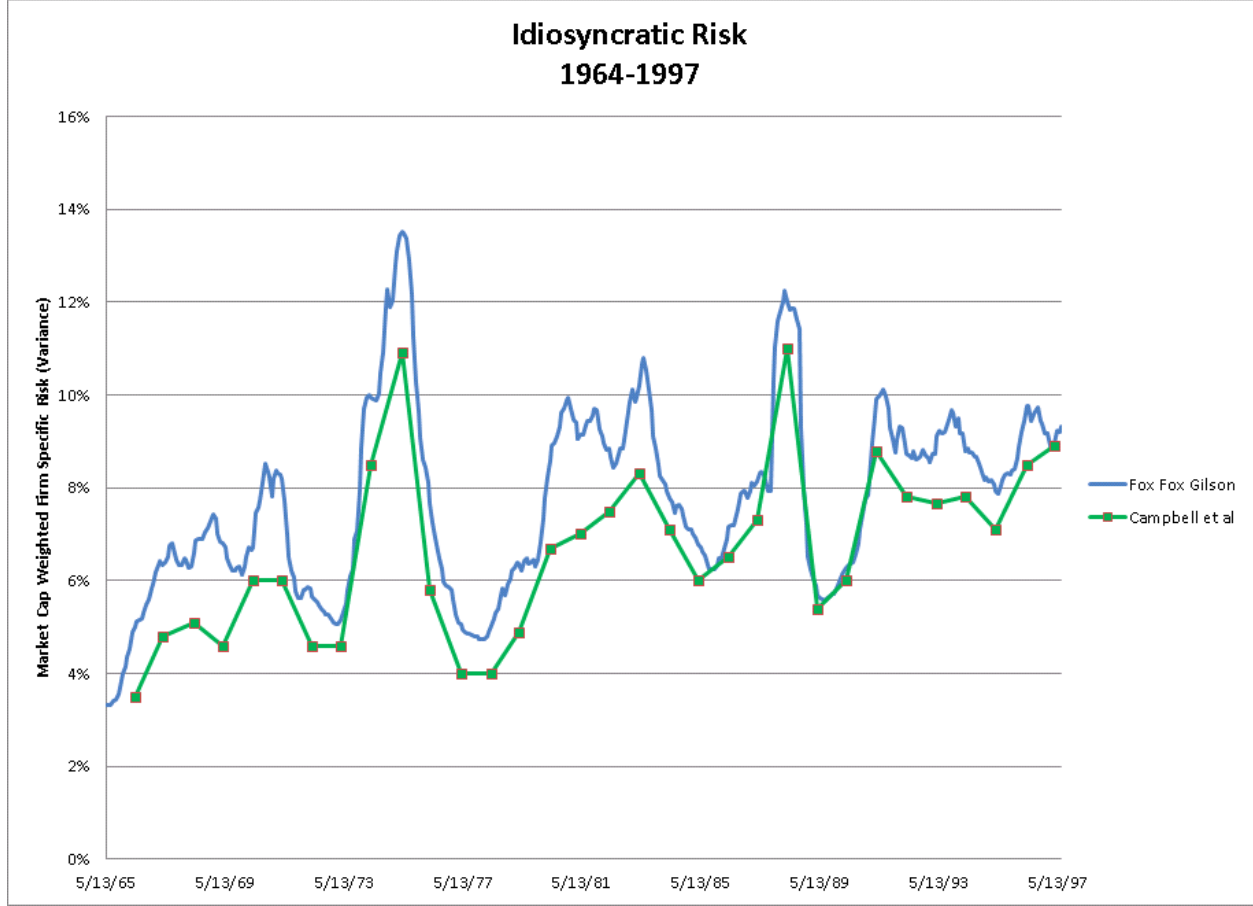


Comparison to Campbell et al.

We use a different estimation method than Campbell et al.³⁷¹ In addition, unlike those authors, our baseline model does not include industry returns. Regardless our results are very similar, particularly since our concern is primarily with the relative changes in idiosyncratic risk rather than with the absolute level:

³⁷¹ Campbell et al., *supra* note 366.

Figure 2A- 2



Industry Returns

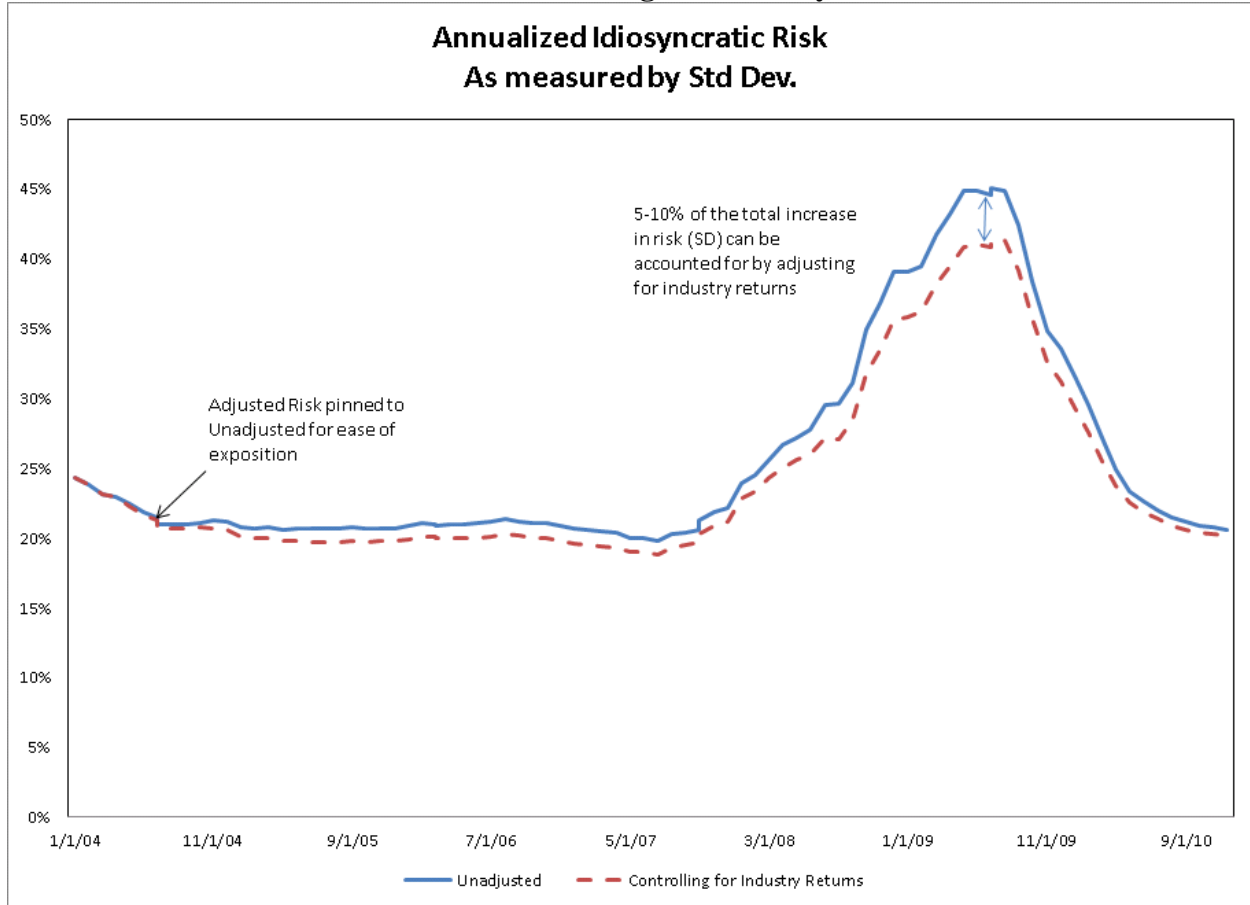
As noted in the body of the text, the most recent crisis hit some industries—financial services and construction—particularly hard. It is therefore conceivable that controlling for industry returns is particularly important during crises. As a robustness check we estimate:

$$Return_{i,t} = \alpha + \beta_i * Return_{mkt,t} + \gamma_i * Return_{ind,t} + \varepsilon_{i,t} \quad \textbf{Equation 2}$$

where $Return_{ind,t}$ is the value-weighted return of firms in the same two-digit SIC code as firm_i.

Including industry returns explains only a small part of the spike in idiosyncratic volatility during the most recent crisis.

Figure 2A- 3
The Effect of Controlling for Industry Returns



Likewise, Table 2A-2 demonstrates that the increase in idiosyncratic risk was not confined to hard-hit industries like finance and construction. Instead, 58 of the 60 sectors we studied saw their risk at least double.

Table 2A- 2:

**Idiosyncratic Risk and Co-movement of Returns with the Broad Market
January 2004 to December 2009**

Two Digit SIC Code	Annualized				Co-Movement of Returns		Change in	Max	
	Description	Idiosyncratic Risk of Firms in ¹		Ratio of Risk 2008-2009 to 2004-2005	("β") with the Broad Market ²		Co-Movement	Number of Firms in SIC Code	
		Each Sector					2008-2009		
		Avg. 2004-2005	Avg. 2008-2009		2004-2005	2008-2009	to 2004-2005		
(1)	(2)	(3)	(4)	(5) (4)/(3)	(6)	(7)	(8) (7)-(6)	(9)	
1)	60	Depository Inst.	2.09%	51.84%	24.75	0.88	1.66	0.78	620
2)	64	Insurance agents, bkrs	4.18%	65.52%	15.67	0.96	1.71	0.75	21
3)	61	Nondeposit. credit inst.	4.25%	49.18%	11.56	0.93	1.42	0.49	63
4)	70	Hotels, other temporary ldg.	5.21%	47.53%	9.13	1.08	1.67	0.59	30
5)	65	Real estate	6.52%	58.32%	8.95	0.89	1.49	0.60	38
6)	63	Insurance carriers	3.58%	31.32%	8.75	0.82	1.18	0.36	174
7)	67	Holding investment offices	2.40%	20.24%	8.43	0.83	1.13	0.31	1559
8)	83	Social services	10.31%	75.34%	7.31	0.89	1.16	0.27	14
9)	62	Security & emdty brokers	4.07%	29.65%	7.28	1.32	1.60	0.28	80
10)	24	Lumber and wood products	3.88%	25.43%	6.56	1.17	1.44	0.27	12
11)	79	Recreational services	6.83%	36.53%	5.35	1.12	1.26	0.14	46
12)	46	Pipelines, except natural gas	2.84%	15.02%	5.28	0.46	0.76	0.31	14
13)	44	Water transportation	6.66%	34.39%	5.17	1.11	1.38	0.27	55
14)	45	Transportation by air	8.24%	42.44%	5.15	1.26	1.14	-0.12	39
15)	26	Paper and allied products	4.11%	21.14%	5.14	1.03	0.91	-0.12	45
16)	27	Printing and publishing	3.70%	18.79%	5.08	0.75	1.23	0.48	66
17)	78	Motion pictures	4.41%	22.18%	5.03	0.81	1.04	0.22	23
18)	15	Gnrl building contractors	8.46%	39.36%	4.65	1.78	1.39	-0.39	31
19)	12	Coal mining	11.58%	53.90%	4.65	1.70	1.82	0.12	17
20)	32	Stone glass & concrete pdcts	8.76%	40.10%	4.58	1.56	1.44	-0.13	28
21)	10	Metal mining	12.87%	57.75%	4.49	1.32	1.14	-0.18	98
22)	55	Car dealers and gas stations	6.38%	28.31%	4.44	1.06	0.96	-0.10	23
23)	37	Transportation equipment	4.35%	19.06%	4.38	1.10	1.07	-0.02	99
24)	22	Textile mill products	9.08%	39.40%	4.34	1.14	1.10	-0.04	17
25)	23	Apparel and other textile	6.49%	27.56%	4.25	1.03	1.09	0.06	33
26)	25	Furniture and fixtures	7.10%	28.14%	3.97	1.01	1.23	0.22	26
27)	48	Communications	4.87%	18.96%	3.89	0.85	0.98	0.13	209
28)	20	Food and kindred products	2.79%	10.31%	3.69	0.61	0.60	-0.02	114
29)	30	Rubber and misc. plastics	6.18%	22.62%	3.66	1.00	1.09	0.09	52
30)	31	Leather and leather products	8.87%	32.20%	3.63	1.34	1.03	-0.32	22
31)	49	Electric, gas, & sanitary svc	3.40%	11.72%	3.45	0.83	0.77	-0.07	167
32)	56	Apparel and accessory	8.32%	27.88%	3.35	1.26	1.07	-0.18	46
33)	57	Furniture & equipment stores	8.14%	26.49%	3.25	1.34	1.07	-0.27	26
34)	72	Personal services	8.09%	24.90%	3.08	0.94	1.11	0.17	22
35)	51	Wholesale (non-durable)	6.70%	20.57%	3.07	0.87	0.85	-0.02	69
36)	40	Railroad Transportation	4.03%	12.00%	2.97	1.19	1.01	-0.18	13
37)	80	Health services	7.30%	21.16%	2.90	0.81	0.77	-0.03	91
38)	39	Misc. mfg industries	8.39%	23.66%	2.82	1.07	1.03	-0.04	50
39)	36	Electrical & electronic equip	6.89%	19.35%	2.81	1.26	1.07	-0.20	446
40)	13	Oil and gas extraction	9.22%	25.74%	2.79	1.31	1.41	0.10	210
41)	53	General merchandise stores	3.87%	10.77%	2.78	0.89	0.63	-0.26	33
42)	16	Heavy constret'n contractors	11.96%	33.09%	2.77	1.60	1.47	-0.14	16
43)	21	Tobacco manufactures Motor freight transit and	3.46%	9.41%	2.72	0.75	0.61	-0.14	10
44)	42	warehousing	4.76%	12.89%	2.71	1.05	0.88	-0.17	35
45)	99	Unclassified	10.67%	28.87%	2.71	0.48	0.37	-0.11	81
46)	59	Miscellaneous retail	6.66%	17.64%	2.65	1.01	0.82	-0.20	92
47)	28	Chemicals and allied products	6.21%	16.36%	2.63	0.87	0.72	-0.15	444
48)	54	Food stores	6.06%	15.86%	2.62	0.92	0.66	-0.26	24
49)	34	Fabricated metal products	6.58%	16.89%	2.57	1.11	1.14	0.03	57
50)	33	Primary metal industries	11.09%	28.41%	2.56	1.75	1.69	-0.06	65

Idiosyncratic Risk and Co-movement of Returns with the Broad Market
January 2004 to December 2009

Two Digit SIC Code	Description	Annualized Idiosyncratic Risk of Firms in ¹ Each Sector		Ratio of Risk 2008-2009 to 2004-2005 (5)/(3)	Co-Movement of Returns ("β") with the Broad Market ²		Change in Co-Movement 2008-2009 to 2004-2005 (8)-(7)	Max Number of Firms in SIC Code (9)
		Avg. 2004-2005	Avg. 2008-2009		2004-2005	2008-2009		
		(3)	(4)		(6)	(7)		
51)	58 Eating and drinking places	5.33%	13.41%	2.52	0.97	0.74	-0.23	68
52)	87 Engineering and management	11.34%	28.50%	2.51	1.18	0.99	-0.19	149
53)	35 Industrial machinery	6.74%	16.85%	2.50	1.18	1.04	-0.15	253
54)	82 Educational services	11.60%	28.83%	2.49	0.94	0.55	-0.39	26
55)	29 Petroleum and coal pdcts	4.06%	9.77%	2.41	1.15	1.04	-0.11	28
56)	50 Wholesale (durables)	8.23%	19.74%	2.40	1.14	0.84	-0.30	132
57)	38 Instruments	7.12%	15.10%	2.12	0.95	0.73	-0.22	295
58)	47 Transportation services	8.00%	16.10%	2.01	1.38	1.03	-0.35	16
59)	73 Business services	8.36%	16.64%	1.99	1.09	0.94	-0.15	599
60)	17 Special trade contractors	14.07%	23.82%	1.69	1.48	1.51	0.03	14

Notes and Sources:

Data obtained from CRSP. The 15 sectors which contain less than 10 firms during the entire 2004-2009 period are not included.

¹ As elsewhere in the paper, the idiosyncratic risk is measured using a regression which removes market effects on the return for each firm. This risk measure is calculated once every month measures the risk over a one year period. Thus for each firm we will have a measure of risk for each firm from January 2004 to January 2005, February 2004 to February 2005, ..., January 2009 to January 2010. The risk for a sector is calculated using market capitalization weights. If firm A has capitalization of \$1.5b and its sector as a whole has an aggregate market capitalization of \$10b, then firm A's idiosyncratic risk will determine 15% of the average idiosyncratic risk of the sector. The risk figures are then averaged over the period listed in columns (3) and (4).

² The measure of market co-movement "β" is a market capitalization weighted average calculated in the same way as the idiosyncratic risk. The interpretation of β is the expected percentage movement in the stock of a firm given a 1% move in the broad market. We calculate a β for each firm and then aggregate all the firms in the sector using a market capitalization weighted average. The β's are then averaged over the period listed in columns (6) and (7).

Three Factor Model

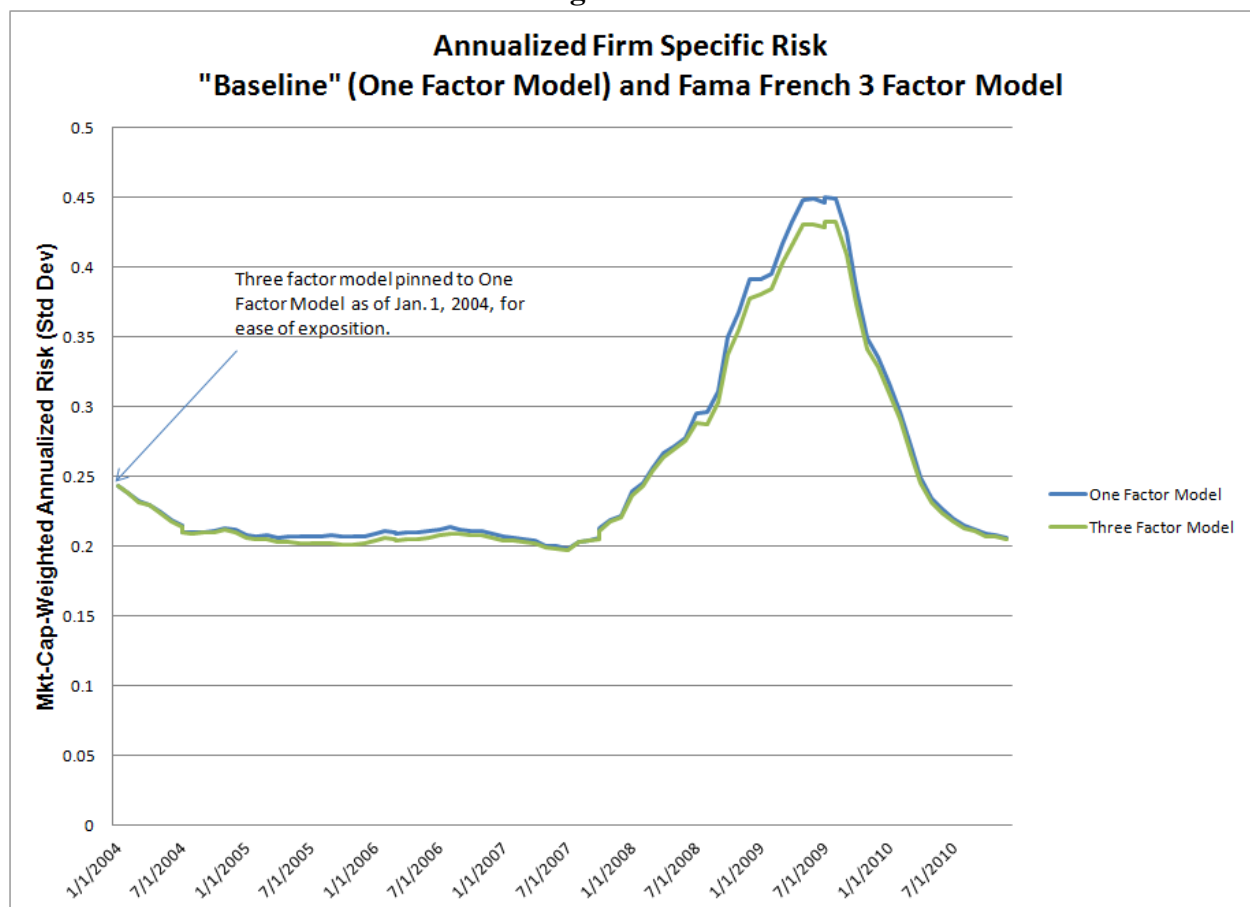
Fama and French have shown that other than a proxy for the market return, two other factors consistently help predict the returns of individual stocks.³⁷² Therefore we estimate their model:

$$Ret_{i,t} = \alpha + \beta_i * Ret_{mkt,t} + \gamma_i * Ret_{SMB,t} + \eta_i * Ret_{HML,t} + \varepsilon_{i,t} \quad \textbf{Equation 3}$$

³⁷² See, e.g., Eugene Fama & Kenneth R. French, *The Capital Asset Pricing Model: Theory and Evidence*, 18 J. ECON. PERSPECTIVES 25 (2004).

where $Return_{SMB,t}$ is the difference between the return on a diversified portfolio composed of small stocks minus that composed of large stocks; and $Return_{HML,t}$ is the return on a portfolio composed of stocks with a high book value to market value minus one composed of stocks with a low book value to market value. The results are very similar to the baseline “one factor” model results:

Figure 2A- 4

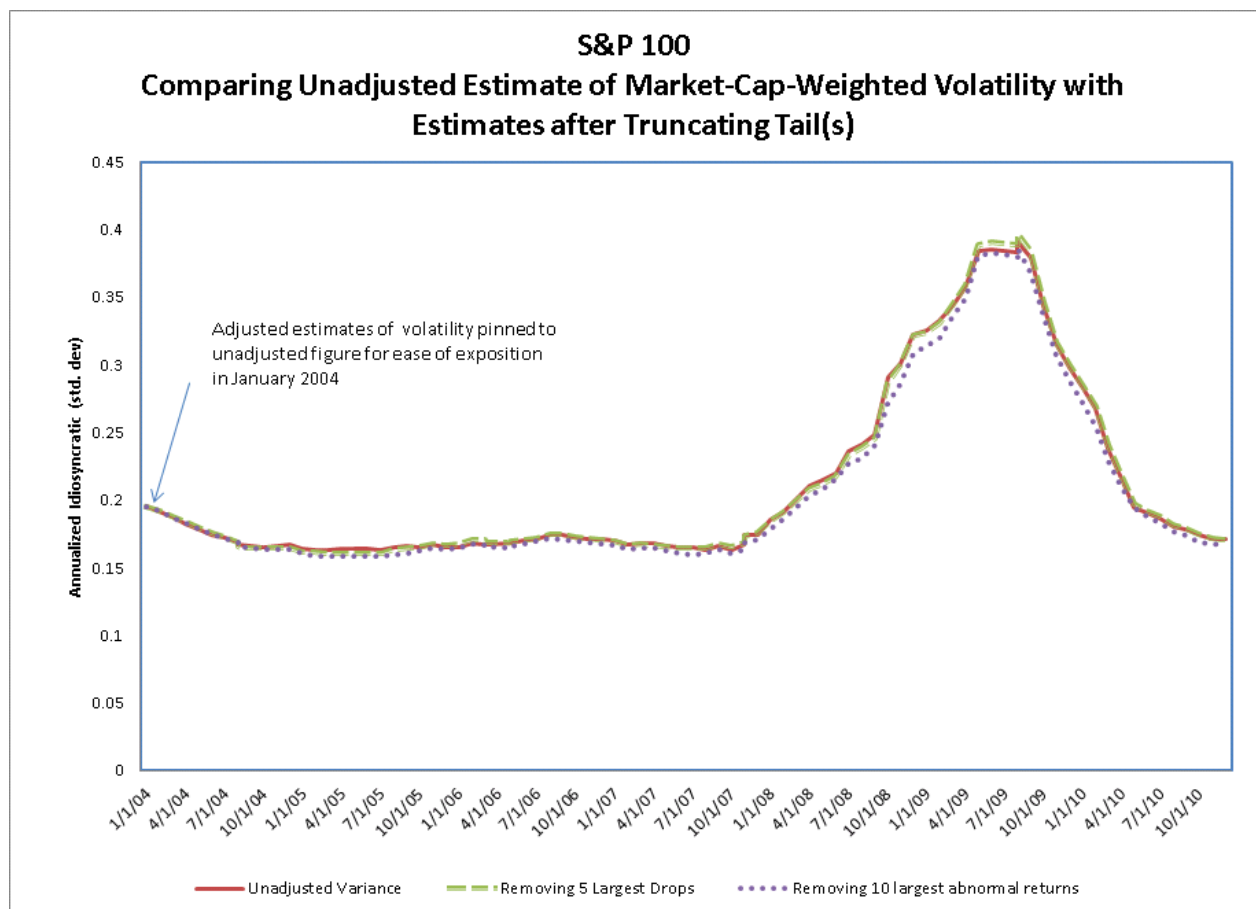


Testing the “Tide-Goes-Out” Hypothesis

Recall that this explanation supposes that the crisis revealed the management of some firms to be incompetent or fraudulent and produced substantial mark-downs in price at those

firms. The remainder of firms would see a modest (relative) mark-up. To test this theory we looked at the portion of volatility explained by the largest net-of-market drops in price. We follow our baseline procedure, except we remove the largest 5 net-of-market drops of each year long regression period from the volatility calculation. If the tide going out was important, the portion explained by these large drops should go up. As shown in the figure below, however, the adjusted volatility looks almost identical. The results are similar if we exclude the 10 largest net-of market returns regardless of sign.

Figure 2A- 5



Share-Price-Drop-Induced Increased Leverage

1. *Controlling for Direct Increase in Risk Stemming from Increases in Leverage.* As shown below, equity value declined substantially during the 2008-2009 crisis increasing leverage:

Table 2A- 3
Leverage Increase 2007-2009
Firms in the S&P 100

	June 2007	March 2009
All Firms		
Market Capitalization	7,631,626,000,000	4,066,529,000,000
Net Debt	3,944,641,400,000	4,592,541,500,000
E/(D+E)	66%	47%
Non-Financial Firms		
Market Capitalization	6,368,326,000,000	3,733,679,000,000
Net Debt	1,134,547,400,000	1,254,592,500,000
E/(D+E)	85%	75%
Financial Firms		
Market Capitalization	1,263,300,000,000	332,850,000,000
Net Debt	2,810,094,000,000	3,337,949,000,000
E/(D+E)	31%	9%

“Net Debt” is short term borrowing plus long term borrowing minus cash and cash equivalents. E= Equity, D= Debt.

Increases in leverage can increase firm-specific risk without a change in the volatility of the market’s implicit valuations of a firm’s future cash flows. This is because equity absorbs (nearly) all changes in the value of the firm’s future cash flows, but the change is concentrated on relatively fewer equity dollars when leverage increases.³⁷³

³⁷³ To illustrate, imagine a firm that will exist for one more period and assume, for purposes of simple exposition, that the risk free rate of interest is zero, there is no systematic risk associated with the issuer, and there is

More formally, if we assume that debt is risk-free,³⁷⁴ then the following formula holds.³⁷⁵

$$\frac{\Delta \text{share price}}{\text{share price}} = \frac{\Delta \text{Equity Value } (\Delta E)}{\text{Equity Value } (E)} = \frac{\Delta \text{Value of Firm's Assets } (\Delta V)}{\text{Value of Firm's Assets } (V)} * \frac{V}{E} = \frac{\Delta V}{V} * \frac{E + \text{Debt } (D)}{E} = \frac{\Delta V}{V} * \left(1 + \frac{D}{E}\right)$$

This implies that

$$\sigma_i^{\text{Unlevered}} = \frac{\text{Equity}_i}{\text{Equity}_i + \text{Debt}_i} * \sigma_i^{\text{Levered}}$$

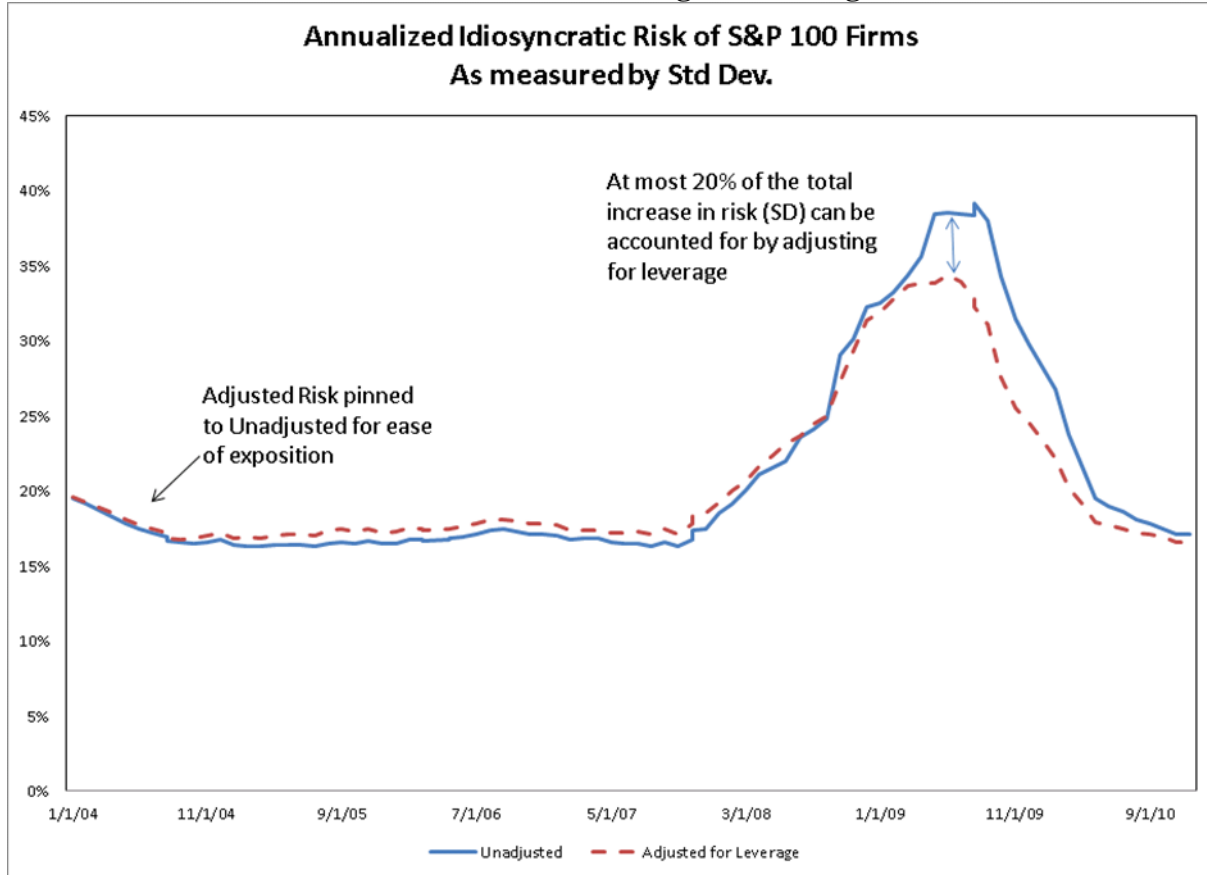
where $\sigma_i^{\text{Unlevered}}$ represents the volatility of stock prices adjusted for leverage. Two firms with identical assets, but different levels of leverage will thus have the same $\sigma_i^{\text{Unlevered}}$ even though their $\sigma_i^{\text{Levered}}$ will be different. Our baseline model presents an estimate of $\sigma_i^{\text{Levered}}$ but the figure below shows that even when we estimate $\sigma_i^{\text{Unlevered}}$ (using Compustat data on the debt of each firm in the S&P 100) the rise in idiosyncratic risk remains steep:

no possibility of news so bad that there is a threat of insolvency. Now assume that the market determines that the expected value of the cash flow available at the end of the period is \$10 and that the firm is capitalized with \$5 in debt and so its equity is valued at \$5. Suppose, at the end of the period, the underlying cash flow turns out to be \$10.50, i.e., 5% (\$.50) above the expected value of \$10. Debt investors will receive \$5 and equity investors would receive \$5.50, representing a 10% return. If instead severe bad news arrives at the beginning of the period and the market determines that the expected value of the firm's cash flow at the end of the period will drop to \$7.00, equity is now valued at \$2.00 and debt, because there is still no risk of insolvency, at \$5.00. If the underlying cash flow turns out to be \$7.35, i.e., also 5% above the expected value of \$7.00, debt investors will receive \$5 and equity investors will receive \$2.35, representing a 17.5% return. If the cash flow turned out to be 5% below its expected value, the size of equity's negative return would be similarly magnified. Thus, a reduction in the value of equity can increase the idiosyncratic risk associated with an issuer's equity even if the risk associated with the issuer's underlying cash flow stays the same.

³⁷⁴ Here, this is a conservative assumption because if, as in reality, debt absorbs some of the change in the firm's assets value, then we will be "over-correcting" for leverage, thereby attributing too much of the increase in the volatility of shares to leverage.

³⁷⁵ See also Stephen Figlewski & Xiaozu Wang, *Is the "Leverage Effect" a Leverage Effect?*, (Working Paper 2000), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=256109.

Figure 2A- 6
The Effect of Controlling for Leverage



2. *Controlling for β Instability Resulting from Leverage Changes.* The equity-value-drop-induced increase in leverage in times of crisis creates a second concern with our results, one related to our estimates of β . Our results are reported in terms of the amount of idiosyncratic volatility, as measured by standard deviation in an issuer's daily net-of-market-returns over the preceding 12 months. Our measure of volatility is thus dependent on "correctly" netting out the market influence. If β is either over or under estimated relative to the "true" relationship between the firm and the market, this will result in an increase in our measure of idiosyncratic risk. Recall that if the estimated β is too large then when the market goes up the model will predict too positive a return, and when the market goes down the model will predict too negative a return. Similarly if the estimated β is too small then when the market goes up, the model will

not predict a large enough gain, and when the market drops the model will predict too small a loss.

For the typical firm, our estimate of β covering the full preceding year will be less accurate during crisis times than in normal times, which biases upward our estimate of firm-specific risk. This is because the true β over the period of the year is likely to change as the firm's leverage changes, and these changes happen more rapidly in crises. In the same way that leverage directly magnifies the idiosyncratic risk associated with the issuer as illustrated above, leverage similarly directly magnifies the systematic riskiness of an issuer's equity (represented by its β times the volatility of the market portfolio) relative to the systematic riskiness of the firm's underlying cash flow.³⁷⁶ Thus this crisis-price-drop-induced change in the β of a firm's equity will, over a twelve-month crisis period, lead to greater inaccuracy in the estimate of the β of its stock than in the case of a twelve-month period in normal times when share prices changes less. This greater inaccuracy in β leads to an increase in our measure of its idiosyncratic risk even if the idiosyncratic risk associated with its underlying cash flow remains unchanged. This raises the possibility that this second leverage related phenomenon can explain the remainder of increase in idiosyncratic risk that we observe in crisis times.

There is some evidence that in fact β instability is correlated with increases in our measure of risk. In Table 2A-4 we test whether there is a relationship between how fast β shifts and increases in our measure of risk. As Table A-4 shows, rapid shifts in β are associated with

³⁷⁶ In the market model, β is the measure of an asset's sensitivity to systematic factors *relative* to the sensitivity of the market as a whole (tautologically the sensitivity of the market as a whole to systemic factors is 1). Since the "market" is really a portfolio of firms, the β s of those firms must at all times average to 1. Thus, all other factors being held constant, whatever happens to firm debt/equity ratios overall, firms whose equity values fall during a crisis more than average, and hence have debt/equity ratios that increase more than average, will have increasing β s during the period. Firms in the opposite position will, *ceteris paribus*, have decreasing β s during the period. Because all the firms in both groups experience changing β s over the period, the single estimate of each such firm's β covering the whole year will be less accurate than in non-crisis years, when their leverage, and hence their β s, stay relatively more stable.

bigger increases in our measure of risk. However, the results are not consistent with β instability driving most of this outcome. For example, firms where the average β changed moderately saw a larger increase in risk than those with large declines in β . This is the opposite of what we would expect if β instability was driving our result.

Table 2A- 4:
Percentage Change in Risk as a Function of Changes in β
Across Industries

Dependent Variable: % Change in Idiosyncratic Risk between 2004-2005 & 2008-2009

Explanatory Variable	Regression Coefficient	T-Statistic
% Change in β btw. 2004 and 2009	13.77	8.33
Large % Decrease ¹ in β btw 2004 & 2009	1.84	2.10
Moderate Change ² in β btw 2004 & 2009	2.38	4.99
Large % Increase ³ in β btw 2004 & 2009	7.86	10.95

Here we use the data represented in Table A-2; that is we calculate risk and β for each firm in CRSP and then take a market cap weighted average of this data across industrial sectors using 2-Digit SIC codes to classify industries. The regressions are then run on these 60 industries.

¹ Dummy variable which is equal to 1 if the industry's firms saw their β 's decrease by more than 20% between 2004-2005 & 2008-2009

² Dummy variable which is equal to 1 if the industry's firms saw their β 's change by less than 20% (either up or down) 2004-2005 & 2008-2009

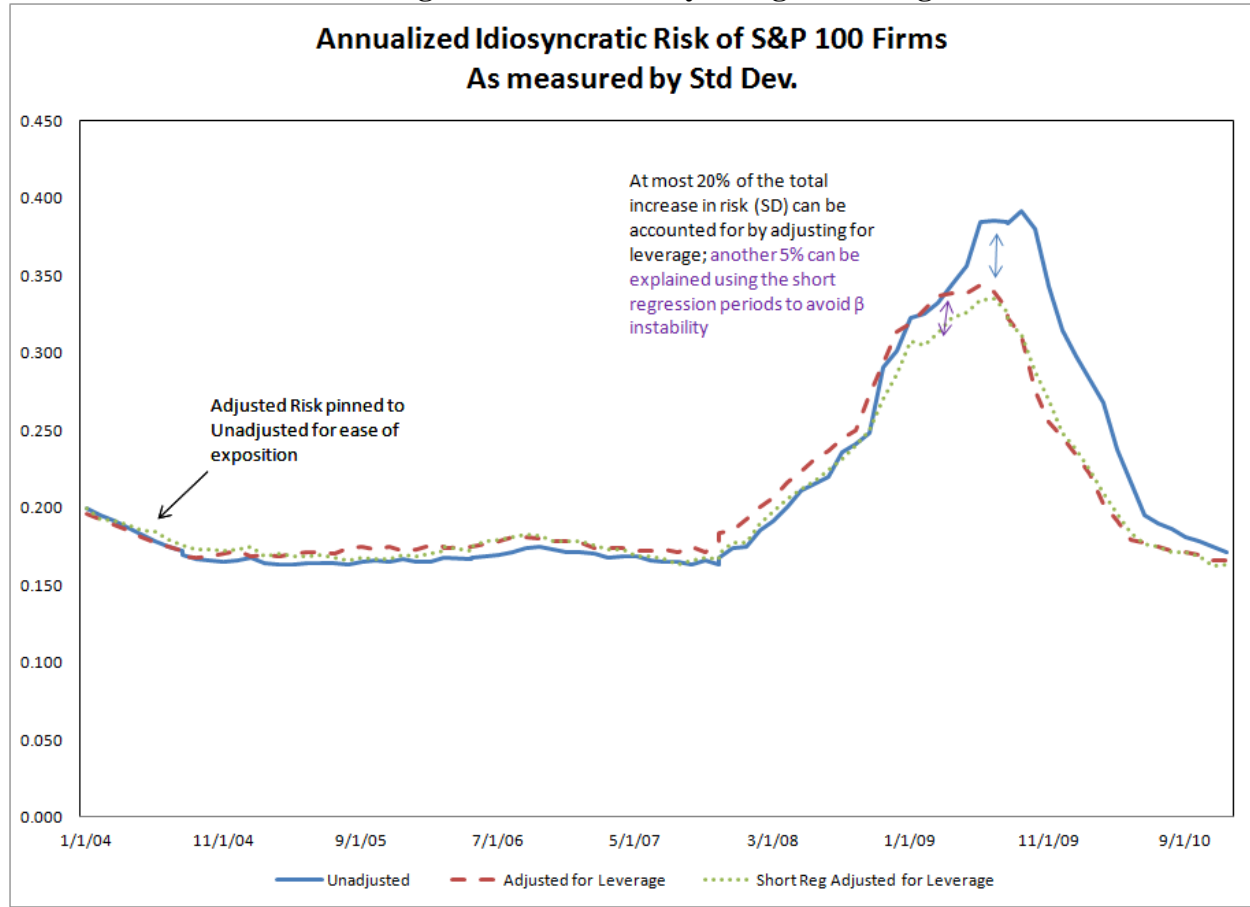
³ Dummy variable which is equal to 1 if the industry's firms saw their β 's increase by more than 20% between 2004-2005 & 2008-2009

Likewise, our indirect control strongly suggests that β instability does not explain much of the remaining increase in idiosyncratic risk that we observe in crisis times. We tested for the effect of this phenomenon by running regression periods using a shorter one-month window, instead of a twelve-month one. The logic is that because the overall price declines associated with the recent financial crisis occurred over a period of several months, leverage (and hence a firm's β) would change in any one month much less than it would over the one-year period over

which we initially ran the regressions underlying the results that we have already reported. This indirect control will not fully correct for β instability because leverage changes can occur within a one month period too and these changes also will be larger during a crisis. But if β instability is driving the increase in idiosyncratic risk, we should see a substantial difference between the one-year regressions (where large changes in leverage over the year would make the linear model fit more poorly) and one-month regressions (where comparatively small changes in leverage would cause the linear model to fit much less poorly). Figure 2A-7 below shows that shortening the regression period does not have a large impact on the percentage increase in idiosyncratic risk.³⁷⁷

³⁷⁷ The long regression periods represent the average standard deviation over the previous year of idiosyncratic returns. In order to make the one-month regression results comparable, we take an average, looking backward, of the results of the current one-month period and the previous 11 one-month periods.

Figure 2A- 7
The Effect of Controlling for Beta Instability Using Short Regression Windows



As an additional way to deal with potential β instability, we also examined the prediction of idiosyncratic volatility imbedded in market traded options. Option prices in part signal the market's expectations about future volatility. Using the Black-Scholes formula on options traded around the strike price, we can back out investors' implicit prediction of the total volatility of a firm. From that we can back out a measure of idiosyncratic volatility. Taking the variance of both sides of equation 1:

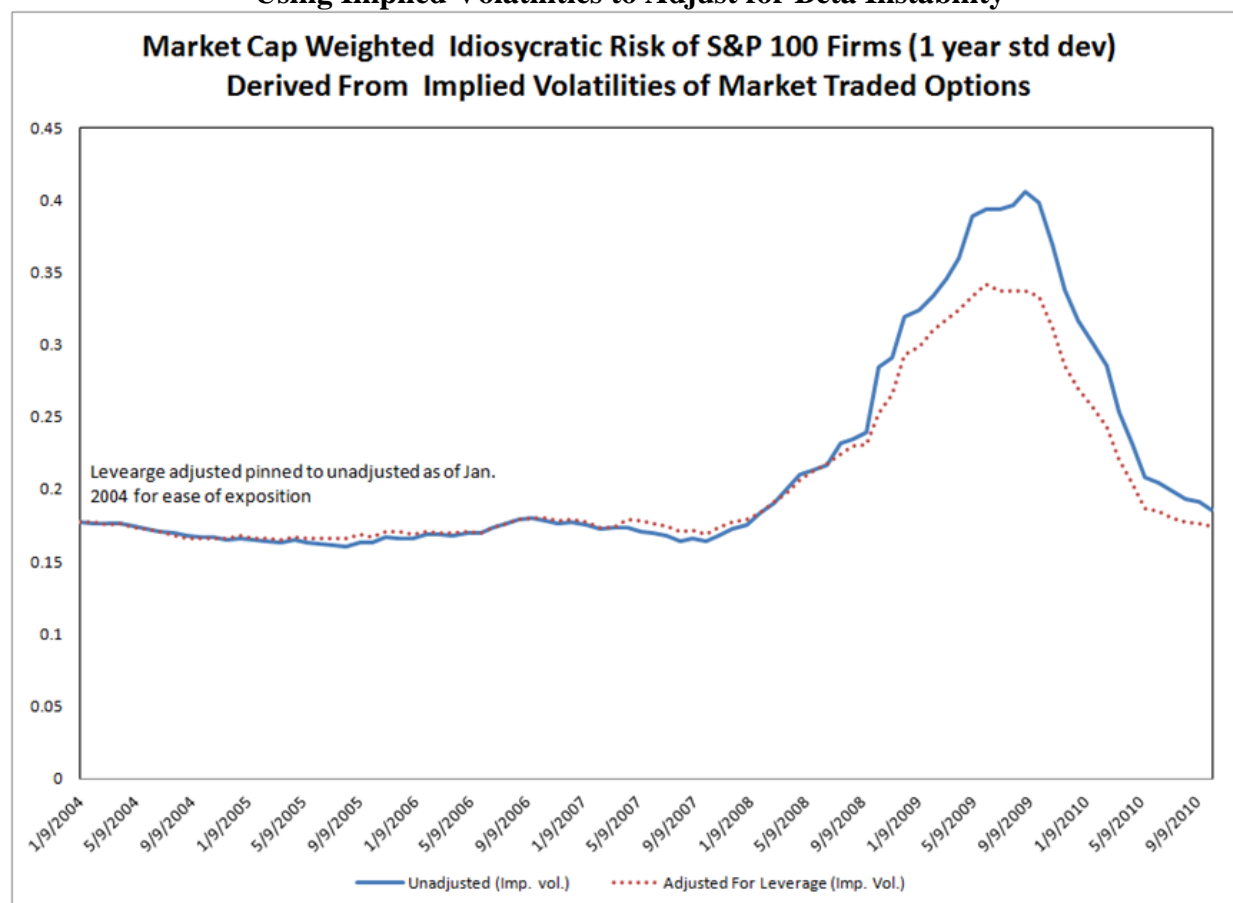
$$Var_{Firm-Total} = \beta_i^2 * Var_{Mkt} + Var_{Firm-Specific} + 2\beta * Covar_{mkt,firm-specific}$$

Note the covariance term is 0 by definition, rearranging we have:

$$Var_{Firm-Specific} = Var_{Firm-Total} - \beta_i^2 * Var_{Mkt}$$

This measure still requires us to estimate β . However, in this case β instability is very likely to lead to a downward bias in our estimate of idiosyncratic risk.³⁷⁸ Nevertheless it still shows a huge spike in idiosyncratic risk even after making the direct leverage adjustment outlined above.

Figure 2A- 8
Using Implied Volatilities to Adjust for Beta Instability



3. *Summary.* When a firm's debt/equity increases, its idiosyncratic volatility increases. Crisis times are usually accompanied by a sharp drop in equity values. This results in notable increases

³⁷⁸We can decompose β for any firm into two parts: our prediction of $\hat{\beta}$ and v where v is an expectations error ($\beta = \hat{\beta} + v$) and $E(v)$ should equal 0 if our estimate is unbiased. This implies that in expectation our $Var_{Firm-Specific}$ will be too small since $E(\hat{\beta}^2) > \beta^2$ (Jensen's inequality: $E(X^2) > [E(X)]^2$). The size of this gap, in expectation, will grow during crises as the variance of v increases (because β is more unstable during crisis). So in expectation this measure of firm-specific risk has a downward bias during crisis (provided that v is independent of the size of the firm).

in most firms' debt/equity ratios, as we saw in 2008-09. However, the sharp drop in share prices in this period turns out to explain only about a quarter of the observed increase in idiosyncratic risk.

An increase in a firm's debt/equity ratio that occurs over time within a single observation period will shift the actual value of its β over this period of time. The market-model-derived estimate of the firm's β for the period will be a less accurate estimate of the actual value of β than when actual β does not shift. Such errors in the estimate of β show up as additional idiosyncratic risk. Thus share-price-drop induced changes in leverage can contribute to our measure of idiosyncratic risk in this way as well. Our indirect control for this problem suggests, however, that it probably does not explain more than an additional 5% of the increase that we witnessed in the 2008-09 period. Moreover, our measure of idiosyncratic risk derived from market-traded options still shows, after adjustment for leverage, a sharp increase in idiosyncratic risk during the recent financial crisis, more than doubling in standard deviation terms, even though this estimate was likely to have been *biased downward* by β instability.

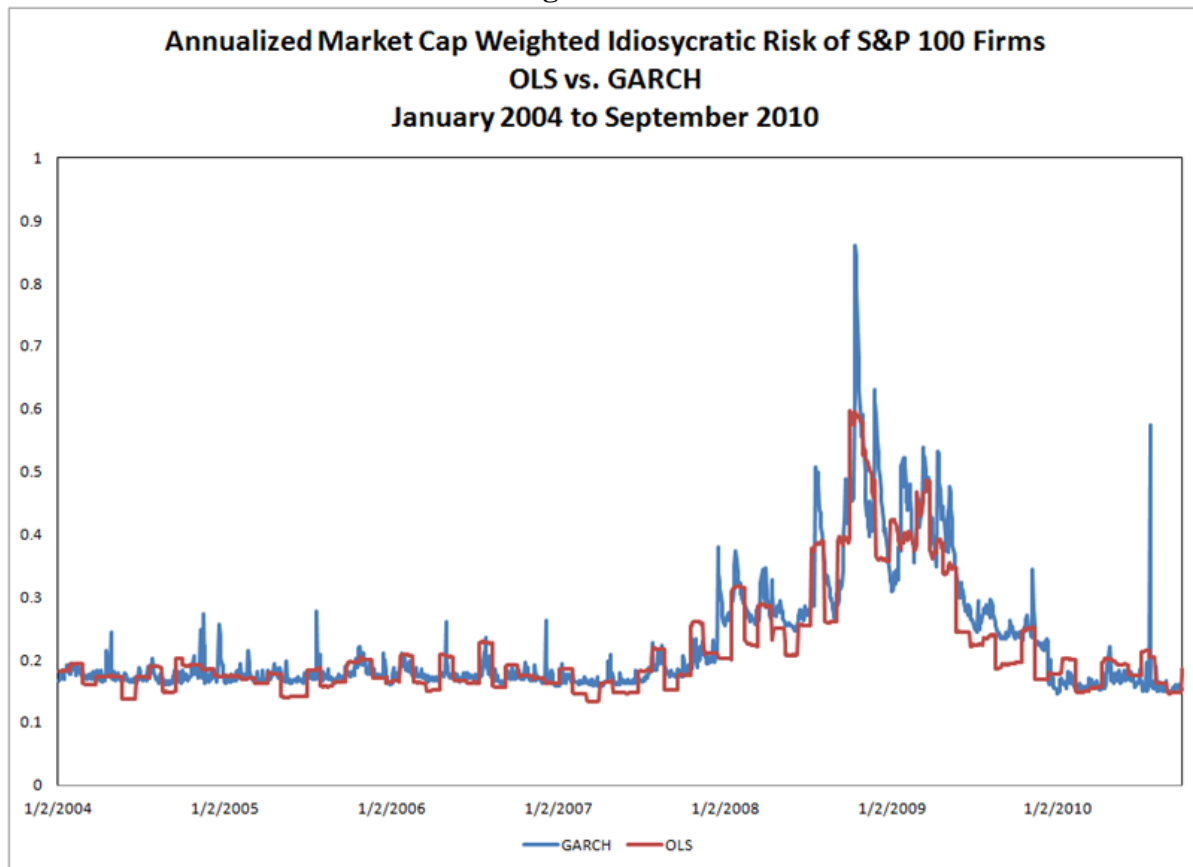
Allowing for a more Complex Error Structure

Along with adding additional controls to the regression and adjusting for the impact of leverage, we also analyzed firm returns assuming the firm-specific volatility can be described within a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) framework. The advantage of GARCH is that we can pick up changes in volatility within the regression period. Under the GARCH(1,1) formulation, the idiosyncratic risk on a given day is assumed to be a function of the yesterday's expected volatility and the size of yesterday's idiosyncratic return.

Because GARCH is estimated via maximum likelihood, relatively long regression windows are needed to ensure convergence.³⁷⁹

Again the results under GARCH look similar to those reached using a simple error. This makes sense intuitively because GARCH estimates of net-of-market returns for each firm will be very similar to the baseline and the difference is only in how to apportion the resulting volatility (our baseline model assumes that the distribution of the error function is constant through time, where GARCH allows it to vary.)

Figure 2A- 9:



³⁷⁹ Therefore GARCH is not particularly well suited to dealing with β instability. The results presented below use 500 day windows for GARCH, but the results for shorter regression windows are very similar.

Appendix 3: Appendix to Does Capital Bear the Corporate Tax After All?

Table 3A- 1:
Comparing a 35% Tax on Normal Rate of Return to Actual Corporate Tax Revenue
For Non-Financial C-Corporations
(Billions of Year-2013 Dollars)

Year	Annual Tax Revenue from C-Corps	Net-Worth of C-Corps	Avg. Risk-Free Rate		Avg. Annual Revenue Raised from 35% Tax on Normal Return		As a % of Actual Revenue	
			Nominal	Real	Nominal	Real	Nominal	Real
			(4)	(5)	Risk-Free Rate	Risk-Free Rate	Risk-Free Rate	Risk-Free Rate
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					=(3)·(4)	=(3)·(5)	=(6)/(2)	=(7)/(2)
Average								
1994 to 2003	164.48	6,538	5.70 %	3.33 %	130.54	76.23	79%	46%
2004 to 2013	206.52	8,110	3.51	1.13	99.54	31.96	48%	15%
1994	157.80	4,601	7.09 %	4.41	114.09	71.02	72%	45%
1995	167.41	4,877	6.57	4.04	112.22	68.89	67%	41%
1996	176.73	5,313	6.44	3.12	119.82	58.04	68%	33%
1997	186.68	5,613	6.35	4.65	124.83	91.39	67%	49%
1998	180.19	6,166	5.26	3.65	113.58	78.80	63%	44%
1999	193.04	6,887	5.65	2.96	136.09	71.38	70%	37%
2000	191.65	8,149	6.03	2.64	171.98	75.39	90%	39%
2001	141.48	8,045	5.02	3.47	141.37	97.68	100%	69%
2002	116.67	7,799	4.61	2.24	125.92	61.04	108%	52%
2003	133.12	7,932	4.01	2.13	111.43	59.25	84%	45%
2004	173.86	7,996	4.27	1.02	119.54	28.42	69%	16%
2005	250.27	8,771	4.29	0.87	131.66	26.81	53%	11%
2006	276.32	8,393	4.80	2.25	140.86	66.22	51%	24%
2007	255.80	8,253	4.63	0.55	133.88	15.98	52%	6%
2008	190.37	7,266	3.66	3.57	93.19	90.86	49%	48%
2009	166.11	7,443	3.26	0.54	85.03	14.14	51%	9%
2010	174.28	8,107	3.22	1.72	91.23	48.79	52%	28%
2011	167.12	7,992	2.78	-0.18	77.81	-5.05	47%	-3%
2012	198.30	8,184	1.80	0.06	51.66	1.79	26%	1%
2013	212.74	8,690	2.35	0.85	71.48	25.80	34%	12%

Notes and Sources: Tax Revenue and Net-Worth of C-Corporations obtained from Table 12 of the IRS Statistics of Income (SOI) Complete Corporations Report. Nominal Risk-Free rate is the average of the 10-year T-bill rate over the course of each year, obtained from the St. Louis Federal Reserve Bank (FRED). The real risk free rate is obtained by removing the Consumer Price Index measure of inflation for that year. Non-Financial industry defined prior to 1998 as all industries except SIC codes 60-67. After 1998, non-financial industries are defined to be as close to equivalent to the pre-1998 definition as possible. This means all industries except NAICS codes 52-53 and 55. However, I count NAICS 532 (rental and leasing services) as non-financial because it was non-financial under the SIC definitions.

Table 3A- 2:
Cashflows of U.S. Non-Financial Corporations
1946-2016

	R+F Base	
	Cashflows as a % of Book Profit	Cashflows as a % of National Income
1946-2016	82.4%	6.6%
1946-1955	65.5%	7.8%
1956-1965	76.6%	7.2%
1966-1975	68.5%	5.9%
1976-1985	82.2%	6.3%
1986-1995	94.8%	5.0%
1996-2005	69.6%	4.3%
2006-2016	119.8%	10.1%

Source:

Federal Reserve Flow of Funds data, 2016. Cashflow = Book profit + Capital Consumption Allowance - Fixed Investment - Net Inventory Investment - Net Financial Investment. See Auerbach (2010).

Table 3A-3

Comparison of Hypothetical R-Base Tax to Actual Corporate Tax: Non-Financial¹ Corporations
Billions of Nominal Dollars

	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970
Expense Equipment and Structures														
(1) Depreciation (+)	15.4	17.1	18.7	20.1	21.5	25.2	27.1	29.8	31.3	34.4	37.5	40.9	45.1	48.6
(2) Net Investment Credits (converted to deduction equivalent) (+) ²	-	-	-	-	-	1.6	2.1	2.7	3.7	4.3	4.4	4.7	3.5	1.6
(3) Net Investment in Equipment and Structures Corporate Non-Financial Firms (BEA) ³	33.4	28.7	30.8	33.4	32.9	35.5	37.3	42.2	50.9	59.2	60.0	65.4	72.9	74.6
(4) Adjust (3) for portion attributable to S-corps ⁴	33.4	28.7	30.8	33.4	32.9	35.5	37.3	42.2	50.9	59.2	60.0	65.4	72.9	74.6
(5) Adjust (4) for coverage diff btw. IRS and NIPA = New Tangible Capital Investment (-) ⁵	30.4	26.5	28.2	30.4	30.1	32.5	33.8	38.0	45.7	53.5	54.5	59.6	66.9	69.2
(6) Net Change Expensing Equipment and Structures [(1)+(2)-(5)]	(15.0)	(9.5)	(9.5)	(10.3)	(8.5)	(5.7)	(4.6)	(5.6)	(10.7)	(14.8)	(12.6)	(14.0)	(18.3)	(18.9)
B. Expense Inventories						-0.15373								
(7) = $\Delta \text{inventory}_{t,t-1}$ (Book value) (-)	3.47	3.87	3.87	3.22	3.22	5.45	6.03	6.84	13.02	14.74	10.23	12.60	19.79	5.77
(8) = $\Delta \text{inventory}_{t,t-1}$ (Tax Accounting) (-) ⁶														
C. Expense Amortizable Capital														
(9) Amortization (+)	2.4	2.0	1.5	1.2	0.9	0.7	0.5	0.3	0.2	0.2	0.2	0.2	0.2	0.7
D. Expense Depletable Capital														
(10) Depletion (+)	3.2	3.1	3.1	3.2	3.4	3.5	3.8	4.2	4.3	4.6	5.1	5.5	5.7	5.4
(11) Change in Base in Shifting to Expensing of Productive Capital (6)-(7)+(9)+(10)	(12.9)	(8.2)	(8.7)	(9.0)	(7.5)	(7.0)	(6.3)	(7.9)	(19.2)	(24.8)	(17.5)	(21.0)	(32.2)	(18.6)
E. Remove Net Financial Income														
(12) Interest Paid (+)	4.5	5.1	5.6	6.1	6.6	7.4	8.1	10.9	9.8	11.8	13.8	16.6	22.0	27.5
(13) Taxable Interest Received (-)	1.3	1.5	1.7	1.8	1.9	2.1	2.5	4.9	2.9	3.5	3.8	4.9	6.6	7.9
(14) Domestic Dividends (-) (adjusted for dividend received deduction) ⁷	0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4	0.4
(15) Net Capital Gains (Adjusted for preferential long term capital gains treatment before TRA '86) (-) ⁸	0.6	0.7	0.7	0.8	1.0	0.9	0.9	1.1	1.1	1.4	1.5	1.6	1.8	1.6
(16) Non-Capital Net Gains (-)	(0.1)	(0.1)	(0.1)	(0.2)	(0.3)	(0.2)	0.0	0.3	0.3	0.3	0.5	0.5	0.7	0.5
(17) Foreign Dividends	0.7	0.8	1.0	1.2	1.4	1.4	1.8	2.4	2.9	2.9	3.1	3.0	3.9	5.0
(18) Net Financial Income (13)+(14)+(15)+(16)+(17)-(12)	(1.8)	(2.8)	(3.0)	(3.4)	(3.8)	(4.2)	(4.3)	(4.3)	(5.1)	(6.2)	(7.7)	(9.2)	(12.5)	(17.1)
F. Tax Base														
(19) Taxable Income Before Credits	36.9	38.8	40.7	39.9	39.1	42.6	47.3	53.0	63.1	69.2	65.8	71.9	71.5	62.2
G. Composition of Productive Capital														
(20) Structures and Equipment ⁹	372	379	393	401	409	422	434	457	488	530	574	630	694	763
(21) Estimated Value of Intangible Capital ¹⁰	70	73	78	81	85	90	95	103	113	125	139	156	176	198
H. Calculations of % of Tax Raised from Normal Return:														
Baseline: Cost of Expensing Productive Capital / Tax Base Net of Fin. Inc. (11)/[(19)-(18)] %	-33.2	-19.7	-20.0	-20.8	-17.4	-14.9	-12.3	-13.8	-28.2	-32.8	-23.8	-25.9	-38.4	-23.4
Leave Depletion and Amortization in Place: [(6)-(7)] / [(19)-(18)] %	-47.7	-32.0	-30.6	-31.1	-27.4	-23.7	-20.6	-21.6	-34.8	-39.2	-31.0	-32.9	-45.4	-31.1
GKS Summary Figure: Net Change in Revenue Switching to R-Base/ Current Base [(11)-[(18)-(17)]]/(19) %	-29.9	-14.0	-14.0	-14.1	-9.4	-6.6	-4.3	-6.8	-22.4	-26.8	-14.9	-16.4	-27.6	-2.3
GKS Summary Figure, Leave Amortization and Depletion in Place [(6)-(7)-[(18)-(17)]]/ (19)%	-45.1	-27.2	-25.4	-25.3	-20.4	-16.3	-13.3	-15.2	-29.5	-33.7	-22.9	-24.3	-35.9	-12.1

Table 3A-3

Comparison of Hypothetical R-Base Tax to Actual Corporate Tax: Non-Financial¹ Corporations
Billions of Nominal Dollars

	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984
Expense Equipment and Structures														
(1) Depreciation (+)	52.6	58.5	64.5	72.2	79.4	86.5	98.9	112.0	127.8	145.9	172.7	197.3	221.8	242.6
(2) Net Investment Credits (converted to deduction equivalent) (+) ²	3.2	6.1	7.6	8.4	12.2	19.4	22.2	25.9	30.4	31.7	39.6	36.1	32.8	35.5
(3) Net Investment in Equipment and Structures Corporate Non-Financial Firms (BEA) ³	76.7	87.6	104.0	117.8	119.2	132.5	159.4	195.3	230.5	249.5	283.0	281.1	271.7	318.7
(4) Adjust (3) for portion attributable to S-corps ⁴	76.7	87.6	104.0	117.8	119.2	132.5	159.4	195.3	230.5	249.5	283.0	281.1	271.7	318.7
(5) Adjust (4) for coverage diff btw. IRS and NIPA = New Tangible Capital Investment (-) ⁵	71.6	82.3	98.8	112.2	113.7	126.0	152.2	185.4	219.0	236.2	271.2	265.5	256.5	302.0
(6) Net Change Expensing Equipment and Structures [(1)+(2)-(5)]	(15.8)	(17.8)	(26.7)	(31.6)	(22.1)	(20.1)	(31.0)	(47.6)	(60.7)	(58.5)	(58.9)	(32.1)	(1.9)	(23.8)
B. Expense Inventories														
(7) = $\Delta \text{inventory}_{t,t-1}$ (Book value) (-)	8.51	23.94	37.91	48.86	5.86	32.38	43.97	45.92	60.91	29.61	48.42	(6.27)	18.26	62.23
(8) = $\Delta \text{inventory}_{t,t-1}$ (Tax Accounting) (-) ⁶														
C. Expense Amortizable Capital														
(9) Amortization (+)	0.7	0.9	0.8	0.7	0.6	0.8	0.8	1.0	1.2	1.2	4.1	3.3	3.5	4.3
D. Expense Depletable Capital														
(10) Depletion (+)	6.0	6.7	9.2	18.8	5.3	5.5	5.6	6.2	7.7	8.7	7.7	6.8	7.4	7.8
(11) Change in Base in Shifting to Expensing of Productive Capital (6)-(7)+(9)+(10)	(17.5)	(34.1)	(54.6)	(60.9)	(22.1)	(46.2)	(68.6)	(86.2)	(112.7)	(78.2)	(95.5)	(15.7)	(9.3)	(74.0)
E. Remove Net Financial Income														
(12) Interest Paid (+)	28.1	30.4	39.1	50.9	51.6	52.6	59.5	73.6	97.0	125.4	162.8	174.1	164.7	188.7
(13) Taxable Interest Received (-)	7.9	8.7	12.5	16.7	16.3	17.9	20.6	26.7	38.5	51.7	72.3	73.4	71.9	84.0
(14) Domestic Dividends (-) (adjusted for dividend received deduction) ⁷	0.4	0.6	1.1	1.9	0.8	1.4	1.6	1.4	1.8	2.2	2.2	2.4	2.5	2.7
(15) Net Capital Gains (Adjusted for preferential long term capital gains treatment before TRA '86) (-) ⁸	1.8	2.3	2.8	3.0	2.7	3.2	4.0	5.2	7.5	8.6	10.1	9.2	9.3	12.1
(16) Non-Capital Net Gains (-)	0.5	0.5	1.5	1.6	2.1	1.8	3.1	3.9	4.8	3.7	4.3	6.5	7.5	11.3
(17) Foreign Dividends	5.9	6.3	8.3	10.8	8.6	13.1	15.4	17.4	24.4	29.4	27.9	26.3	27.4	31.4
(18) Net Financial Income (13)+(14)+(15)+(16)+(17)-(12)	(17.5)	(18.3)	(21.3)	(27.8)	(29.8)	(28.3)	(30.3)	(36.4)	(44.4)	(59.2)	(74.0)	(82.5)	(73.6)	(78.5)
F. Tax Base														
(19) Taxable Income Before Credits	70.7	81.3	101.7	131.4	133.4	166.9	191.2	212.7	251.1	222.3	219.6	185.2	196.2	231.8
G. Composition of Productive Capital														
(20) Structures and Equipment⁹	835	904	1,019	1,251	1,388	1,524	1,694	1,922	2,213	2,538	2,867	3,034	3,101	3,266
(21) Estimated Value of Intangible Capital¹⁰	222	240	278	350	398	448	510	593	700	821	950	1,028	1,075	1,158
H. Calculations of % of Tax Raised from Normal Return:														
Baseline: Cost of Expensing Productive Capital / Tax Base Net of Fin. Inc. (11)/[(19)-(18)] %	-19.9	-34.3	-44.4	-38.3	-13.5	-23.6	-31.0	-34.6	-38.1	-27.8	-32.5	-5.9	-3.4	-23.8
Leave Depletion and Amortization in Place: [(6)-(7)] / [(19)-(18)] %	-27.5	-41.9	-52.6	-50.5	-17.1	-26.9	-33.9	-37.5	-41.1	-31.3	-36.5	-9.6	-7.5	-27.7
GKS Summary Figure: Net Change in Revenue Switching to R-Base/ Current Base [(11)-[(18)-(17)]]/(19) %	0.0	-19.5	-32.8	-25.2	5.8	-10.7	-20.1	-23.4	-27.2	-8.5	-9.8	36.1	32.8	2.0
GKS Summary Figure, Leave Amortization and Depletion in Place [(6)-(7)-[(18)-(17)]]/ (19)%	-9.6	-28.8	-42.6	-40.1	1.4	-14.5	-23.4	-26.8	-30.8	-13.0	-15.1	30.6	27.2	-3.3

Table 3A-3

Comparison of Hypothetical R-Base Tax to Actual Corporate Tax: Non-Financial¹ Corporations
Billions of Nominal Dollars

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Expense Equipment and Structures													
(1) Depreciation (+)	277.5	271.8	265.3	268.3	274.4	269.9	270.9	281.9	294.6	325.7	350.5	380.3	407.1
(2) Net Investment Credits (converted to deduction equivalent) (+) ²	39.7	19.1	6.8	(1.5)	(1.0)	(0.3)	(0.2)	(0.1)	(0.1)	-	-	-	-
(3) Net Investment in Equipment and Structures Corporate Non-Financial Firms (BEA) ³	337.3	319.9	316.6	337.2	361.6	389.4	374.4	385.8	422.9	465.9	529.7	570.1	613.1
(4) Adjust (3) for portion attributable to S-corps ⁴	337.3	300.4	292.5	307.3	326.1	348.6	334.6	343.5	375.5	414.3	468.9	504.6	541.8
(5) Adjust (4) for coverage diff btw. IRS and NIPA = New Tangible Capital Investment (-) ⁵	319.2	282.8	275.5	289.6	306.9	327.7	314.7	324.4	355.3	395.2	450.4	486.8	523.7
(6) Net Change Expensing Equipment and Structures [(1)+(2)-(5)]	(2.0)	8.1	(3.4)	(22.8)	(33.4)	(58.1)	(44.1)	(42.5)	(60.8)	(69.5)	(99.9)	(106.5)	(116.6)
B. Expense Inventories													
(7) = $\Delta \text{inventory}_{t,t-1}$ (Book value) (-)	39.94	(24.22)	12.87	21.25	18.57	5.45	(7.30)	26.00	26.51	49.31	19.72	26.61	18.10
(8) = $\Delta \text{inventory}_{t,t-1}$ (Tax Accounting) (-) ⁶											49.44	26.29	48.77
C. Expense Amortizable Capital													
(9) Amortization (+)	5.1	10.6	12.7	15.1	19.5	23.9	27.5	27.5	27.2	28.7	33.7	38.3	45.0
D. Expense Depletable Capital													
(10) Depletion (+)	7.5	8.7	8.1	8.3	8.8	9.2	8.1	8.5	8.0	9.0	9.6	9.7	9.6
(11) Change in Base in Shifting to Expensing of Productive Capital (6)-(7)+(9)+(10)	(29.4)	51.6	4.5	(20.7)	(23.7)	(30.5)	(1.2)	(32.6)	(52.1)	(81.1)	(76.3)	(85.1)	(80.1)
E. Remove Net Financial Income													
(12) Interest Paid (+)	205.6	214.3	214.3	241.6	289.5	296.8	287.7	252.2	235.2	250.4	286.4	305.9	337.5
(13) Taxable Interest Received (-)	94.3	92.3	87.9	99.5	119.2	146.8	148.1	135.7	127.5	134.2	155.3	176.4	195.4
(14) Domestic Dividends (-) (adjusted for dividend received deduction) ⁷	1.8	1.4	1.2	1.3	1.2	1.3	1.3	1.3	1.4	1.2	3.0	1.3	1.4
(15) Net Capital Gains (Adjusted for preferential long term capital gains treatment before TRA '86) (-) ⁸	14.7	22.6	36.3	38.3	44.8	30.1	27.1	23.8	27.3	34.7	37.7	41.7	63.0
(16) Non-Capital Net Gains (-)	12.9	11.1	11.0	20.9	16.5	9.8	11.3	9.6	10.7	14.9	18.3	24.7	27.4
(17) Foreign Dividends	39.3	37.8	40.8	65.8	58.2	64.4	56.0	50.5	54.0	59.5	69.5	86.0	96.2
(18) Net Financial Income (13)+(14)+(15)+(16)+(17)-(12)	(81.8)	(86.9)	(77.9)	(81.6)	(107.9)	(108.9)	(100.0)	(81.8)	(68.3)	(65.6)	(72.1)	(61.8)	(50.4)
F. Tax Base													
(19) Taxable Income Before Credits	231.3	225.5	260.7	323.3	305.9	297.8	269.0	276.2	312.2	379.0	417.7	472.8	502.4
G. Composition of Productive Capital													
(20) Structures and Equipment⁹	3,422	3,547	3,705	3,921	4,129	4,344	4,427	4,566	4,780	5,035	5,327	5,590	5,891
(21) Estimated Value of Intangible Capital¹⁰	1,241	1,315	1,404	1,518	1,633	1,755	1,826	1,923	2,055	2,210	2,393	2,571	2,772
H. Calculations of % of Tax Raised from Normal Return:													
Baseline: Cost of Expensing Productive Capital / Tax Base Net of Fin. Inc.	-9.4	16.5	1.3	-5.1	-5.7	-7.5	-0.3	-9.1	-13.7	-18.2	-15.6	-15.9	-14.5
(11)/[(19)-(18)] %													
Leave Depletion and Amortization in Place: [(6)-(7)] / [(19)-(18)] %	-13.4	10.3	-4.8	-10.9	-12.6	-15.6	-10.0	-19.1	-22.9	-26.7	-24.4	-24.9	-24.4
GKS Summary Figure: Net Change in Revenue Switching to R-Base/ Current Base [(11)-[(18)-(17)]]/(19) %	22.7	61.4	31.6	18.9	27.5	26.3	36.7	17.8	5.2	-4.1	-1.0	-4.9	-5.9
GKS Summary Figure, Leave Amortization and Depletion in Place [(6)-(7)-[(18)-(17)]]/ (19)%	17.2	52.8	23.6	11.6	18.3	15.2	23.5	4.8	-6.1	-14.0	-11.4	-15.1	-16.8

Table 3A-3

Comparison of Hypothetical R-Base Tax to Actual Corporate Tax: Non-Financial¹ Corporations
Billions of Nominal Dollars

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Expense Equipment and Structures													
(1) Depreciation (+)	426.0	452.4	473.6	497.0	541.7	524.7	523.2	402.9	420.9	446.9	567.7	543.4	558.0
(2) Net Investment Credits (converted to deduction equivalent) (+) ²	-	-	-	-	-	-	-	-	-	-	-	-	-
(3) Net Investment in Equipment and Structures Corporate Non-Financial Firms (BEA) ³	648.7	689.5	753.4	713.5	617.2	621.8	664.2	758.1	860.6	929.8	938.0	722.1	748.0
(4) Adjust (3) for portion attributable to S-corps ⁴	572.8	605.6	658.9	620.6	531.4	528.4	559.4	637.8	711.1	765.5	771.4	607.6	633.1
(5) Adjust (4) for coverage diff btw. IRS and NIPA = New Tangible Capital Investment (-) ⁵	555.2	587.3	639.5	602.1	515.9	512.4	540.9	608.9	677.4	727.0	735.9	574.9	596.9
(6) Net Change Expensing Equipment and Structures [(1)+(2)-(5)]	(129.2)	(134.8)	(165.9)	(105.2)	25.8	12.4	(17.7)	(206.0)	(256.4)	(280.2)	(168.2)	(31.5)	(38.9)
B. Expense Inventories													
(7) = $\Delta \text{inventory}_{t,t-1}$ (Book value) (-)	24.62	38.74	52.69	(59.36)	(13.66)	30.29	77.16	82.90	67.73	33.41	(29.36)	(79.05)	49.28
(8) = $\Delta \text{inventory}_{t,t-1}$ (Tax Accounting) (-) ⁶	39.60	54.51	100.57	-38.02	17.10	38.12	88.32	72.65	84.55	39.13	-23.43	-116.64	74.19
C. Expense Amortizable Capital													
(9) Amortization (+)	54.9	63.2	72.2	81.3	85.9	88.6	98.4	109.8	112.5	124.2	138.4	146.3	145.5
D. Expense Depletable Capital													
(10) Depletion (+)	9.0	9.0	9.4	9.2	8.9	9.6	11.3	14.7	14.3	18.2	20.5	20.7	21.7
(11) Change in Base in Shifting to Expensing of Productive Capital (6)-(7)+(9)+(10)	(89.9)	(101.4)	(136.9)	44.7	134.2	80.3	14.8	(164.3)	(197.4)	(171.2)	20.0	214.5	79.0
E. Remove Net Financial Income													
(12) Interest Paid (+)	356.3	388.0	457.0	443.8	396.2	373.3	374.3	428.1	495.1	573.0	525.1	442.3	416.2
(13) Taxable Interest Received (-)	194.0	207.0	248.5	233.6	209.5	197.5	207.5	248.2	303.2	332.6	259.4	179.4	145.5
(14) Domestic Dividends (-) (adjusted for dividend received deduction) ⁷	1.3	1.4	1.5	1.2	1.0	1.2	1.3	1.6	2.3	2.2	2.6	1.9	3.2
(15) Net Capital Gains (Adjusted for preferential long term capital gains treatment before TRA '86) (-) ⁸	69.2	104.1	121.5	80.0	52.0	56.6	63.0	80.7	92.3	105.4	68.7	41.1	56.4
(16) Non-Capital Net Gains (-)	31.8	26.8	27.2	2.0	(5.0)	(1.9)	10.0	15.8	37.9	32.8	(6.2)	(2.4)	13.4
(17) Foreign Dividends	92.2	107.3	110.0	98.4	93.3	107.6	122.2	437.8	145.2	167.8	221.9	239.1	242.6
(18) Net Financial Income (13)+(14)+(15)+(16)+(17)-(12)	(60.0)	(48.7)	(58.2)	(127.0)	(138.8)	(119.9)	(92.5)	(81.7)	(59.4)	(99.9)	(200.6)	(222.3)	(197.7)
F. Tax Base													
(19) Taxable Income Before Credits	485.7	515.6	553.8	438.0	390.2	452.0	574.7	854.0	909.3	909.7	783.6	704.5	805.4
G. Composition of Productive Capital													
(20) Structures and Equipment⁹	6,178	6,489	6,917	7,262	7,499	7,720	8,390	9,178	9,940	10,467	11,141	10,691	10,974
(21) Estimated Value of Intangible Capital¹⁰	2,975	3,196	3,484	3,741	3,950	4,157	4,618	5,163	5,713	6,147	6,709	6,246	6,449
H. Calculations of % of Tax Raised from Normal Return:													
Baseline: Cost of Expensing Productive Capital / Tax Base Net of Fin. Inc. (11)/[(19)-(18)] %	-16.5	-18.0	-22.4	7.9	25.4	14.0	2.2	-17.6	-20.4	-17.0	2.0	23.1	7.9
Leave Depletion and Amortization in Place: [(6)-(7)] / [(19)-(18)] %	-28.2	-30.8	-35.7	-8.1	7.5	-3.1	-14.2	-30.9	-33.5	-31.1	-14.1	5.1	-8.8
GKS Summary Figure: Net Change in Revenue Switching to R-Base/ Current Base [(11)-[(18)-(17)]]/(19) %	-6.2	-10.2	-14.2	39.2	70.0	44.3	18.7	-9.7	-15.2	-7.8	28.2	62.0	34.4
GKS Summary Figure, Leave Amortization and Depletion in Place [(6)-(7)-[(18)-(17)]]/ (19)%	-19.3	-24.2	-29.0	18.5	45.7	22.6	-0.4	-24.3	-29.1	-23.5	7.9	38.3	13.6

Table 3A-3

Comparison of Hypothetical R-Base Tax to Actual Corporate Tax: Non-Financial¹ Corporations
Billions of Nominal Dollars

	2011	2012	2013
Expense Equipment and Structures			
(1) Depreciation (+)	675.5	555.0	560.1
(2) Net Investment Credits (converted to deduction equivalent) (+) ²	-	-	-
(3) Net Investment in Equipment and Structures Corporate Non-Financial Firms (BEA) ³	852.2	961.4	996.7
(4) Adjust (3) for portion attributable to S-corps ⁴	719.2	824.2	844.6
(5) Adjust (4) for coverage diff btw. IRS and NIPA = New Tangible Capital Investment (-) ⁵	680.5	764.5	778.4
(6) Net Change Expensing Equipment and Structures [(1)+(2)-(5)]	(5.0)	(209.5)	(218.3)
B. Expense Inventories			
(7) = $\Delta \text{inventory}_{t,t-1}$ (Book value) (-)	84.63	69.17	56.09
(8) = $\Delta \text{inventory}_{t,t-1}$ (Tax Accounting) (-) ⁶	91.06	71.60	56.60
C. Expense Amortizable Capital			
(9) Amortization (+)	149.9	153.4	160.6
D. Expense Depletable Capital			
(10) Depletion (+)	25.0	26.1	25.4
(11) Change in Base in Shifting to Expensing of Productive Capital	85.4	(99.2)	(88.4)
(6)-(7)+(9)+(10)			
E. Remove Net Financial Income			
(12) Interest Paid (+)	413.4	415.2	401.6
(13) Taxable Interest Received (-)	146.9	133.6	121.5
(14) Domestic Dividends (-) (adjusted for dividend received deduction) ⁷	1.6	1.5	1.7
(15) Net Capital Gains (Adjusted for preferential long term capital gains treatment before TRA '86) (-) ⁸	60.2	76.5	68.0
(16) Non-Capital Net Gains (-)	20.5	31.7	19.6
(17) Foreign Dividends	192.5	201.0	210.6
(18) Net Financial Income (13)+(14)+(15)+(16)+(17)-(12)	(184.2)	(171.9)	(190.8)
F. Tax Base			
(19) Taxable Income Before Credits	768.2	879.2	922.5
G. Composition of Productive Capital			
(20) Structures and Equipment⁹	11,430	11,814	12,218
(21) Estimated Value of Intangible Capital¹⁰	6,815	6,930	7,050
H. Calculations of % of Tax Raised from Normal Return:			
Baseline: Cost of Expensing Productive Capital / Tax Base Net of Fin. Inc.	9.0	-9.4	-7.9
(11)/[(19)-(18)] %			
Leave Depletion and Amortization in Place: [(6)-(7)] / [(19)-(18)] %	-9.4	-26.5	-24.6
GKS Summary Figure: Net Change in Revenue Switching to R-Base/ Current Base [(11)-[(18)-(17)]]/(19) %	35.1	8.3	11.1
GKS Summary Figure, Leave Amortization and Depletion in Place [(6)-(7)-[(18)-(17)]]/ (19)%	12.3	-12.1	-9.1

Notes and Sources:

The data come from the yearly IRS Statistics of Income Complete Corporation Report and Corporation Sourcebook, unless otherwise noted.

Prior to 1986, to ease the burden of hand collecting the data, the figures include S-corporations. In the period from the introduction of the S-Corp in 1958 to 1986, the inclusion or exclusion of S-Corps makes a negligible difference because S-Corps made up an economically small portion of total corporate capital etc. (e.g., less than 4% of depreciable capital in 1986 and less in previous years). The data for 1958 and 1960 are interpolated due to difficulties in collecting it.

Grey highlighting indicates that the year contains the trough of an NBER designated recession

- ¹ In 1998 the IRS switched to reporting industry figures by NAICS code, rather than SIC code. I use the SIC based definition of "Finance, insurance, and real estate" throughout the analysis. To convert the NAICS definition to be equivalent to the SIC based definition of finance, I use the following industries: "Finance and insurance," "Real estate and rental and leasing" net of the sub-major-industry "Rental and leasing services," and "Management of Holding Companies." This follows the definition used by Cronin et al. (2013).
- ² Investment credits net of recapture are grossed up into a "deduction" equivalent using the effective tax rate on taxable income for that year (generally about the statutory rate).
- ³ Obtained from BEA Table 4.7: "Investment in Private Nonresidential Fixed Assets by Industry Group and Legal Form of Organization." Note the BEA considers S-Corps to be corporate organizations. The BEA figure also represents "net investment." Investment in *new* structures and equipment by non-financial corporations is larger than the figure displayed above because the net investment figure represents the investment in new structures and equipment reduced by net sales of used equipment and structures by those corporations to individuals and government. Those net sales of used equipment and structures are positive in every year. Under a cash flow tax these sales of used capital goods from corporations to individuals would be fully included by the corporation with no accompanying basis deduction. This effect can be mimicked by using the BEA net investment figure and eliminating all capital gains and non-capital gains. If the figure for new capital goods is employed and all capital gains are eliminated, then the corporate cash flow base will be understated by the amount of the net sales of used equipment and structures to the non-corporate sector.
- ⁴ As explained above, there is no adjustment for S-Corps' investment prior to 1986 because the relevant tax data presented above also includes S-corps. Starting in 1986 S-Corp investment is apportioned based on that year's ratio of S-Corp non-financial depreciation to C-Corp non-financial depreciation.
- ⁵ Uses data from BEA Table 7.13: "Relation of Consumption of Fixed Capital in the National Income and Product Accounts to Depreciation and Amortization as Published by the Internal Revenue Service." The BEA data covers a somewhat different universe than the IRS data. For example, the IRS data includes foreign branches, but BEA does not. Likewise the tax code allows for the immediate deduction of some items the BEA considers to be capital expenditures, like many expenditures on mines or well exploration. The BEA figure is therefore adjusted for coverage differences by scaling the BEA figure by the ratio of IRS:BEA depreciation of *physical capital*.
- ⁶ After 1986, tax inventory charges will likely systematically exceed book inventory charges because the TRA '86 required capitalization of some expenses not traditionally capitalized into inventory in book accounting. The size of this difference is fairly small though (~\$10 B a year from 1995 to 2013). The public IRS data is not sufficient to calculate the tax inventory charges prior to 1995.
- ⁷ The fact that most domestic dividends are already excluded from corporate taxation via the various domestic dividend deductions is adjusted for by taking 20% of the raw domestic dividends as subject to full taxation, following GKS. From 2000-2013 this adjustment was almost exactly correct based on the additional data provided in those years dividends are exempted from taxation.
- ⁸ The preferential rate given to long term capital gains under tax law prior to 1986 (and hence partial exclusion of such gains from the tax base in the pre 1986 period) is adjusted for by reducing long-term capital gains by 20%/the effective tax rate on taxable income for that year.
- ⁹ Data obtained from BEA Table 4.1.
- ¹⁰ Data correspond to estimates of Corrado and Hulten (2010) made using perpetual inventory method. For 2007 to 2013, I tack on the estimates using Peters and Taylor (2017), who use a similar methodology for public companies.